

# Lezioni operative nell'ambito del progetto “INNOVA.RE”

P.O.R. Sardegna 2007/2013  
**Cagliari, 23 novembre 2012**

---

## Ricerca, creare, disseminare le informazioni scientifiche Lezioni operative

Francesca Di Donato  
Università di Pisa  
[didonato@sp.unipi.it](mailto:didonato@sp.unipi.it)

# Programma

## **0. Il Web come ambiente di ricerca**

Storia e architettura del WWW

## **1. Guida all'uso dei motori di ricerca**

1.1. Introduzione alla topologia del Web

1.2. I principali motori di ricerca e le regole generali per cercare informazioni sul web

1.3. Gli strumenti per la ricerca scientifica: motori, basi di dati, etc.

# Programma

## **2. Creare, gestire, condividere bibliografie scientifiche**

2.1. Web2.0 e social web

2.2. Introduzione all'uso di Citeulike, Connotea, Zotero

## **3. Disseminare i propri risultati**

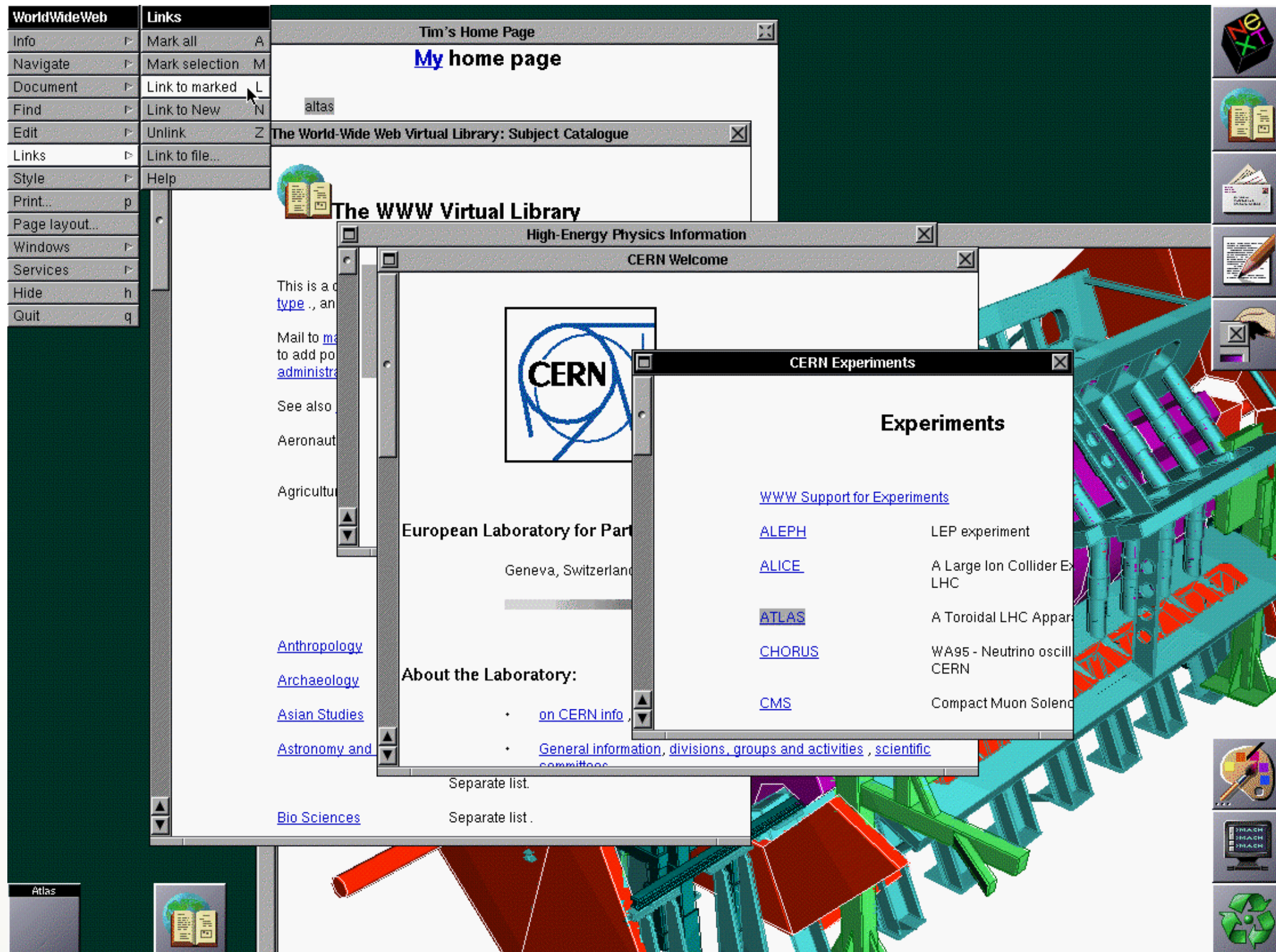
3.1. Il modello di pubblicazione tradizionale

3.2. Il modello di pubblicazione ad accesso aperto

# o. Il Web come ambiente di ricerca



# La nascita del Web



sulle spalle dei giganti...

# Internet

**Architettura**

distribuita

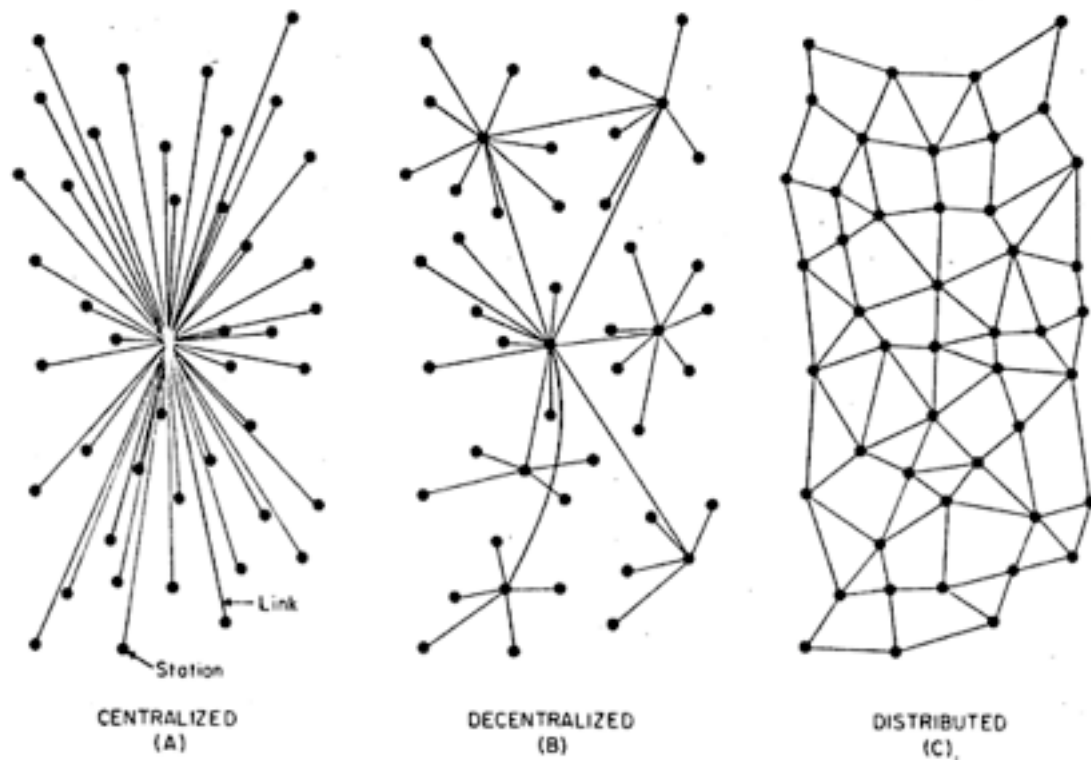


FIG. 1 - Centralized, Decentralized and Distributed Networks

# Internet



## Architettura

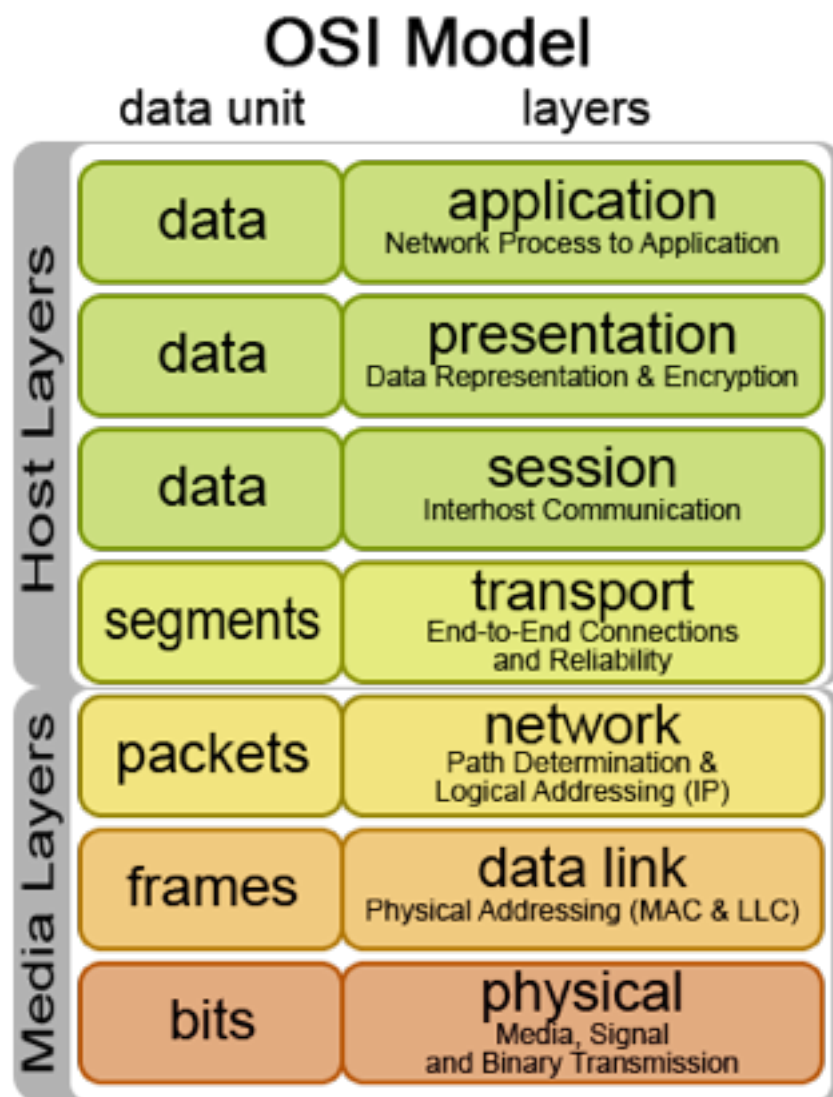
distribuita

flessibile

# Internet

## Architettura

distribuita  
flessibile  
aperta





# Internet



## **Architettura**

distribuita

flessibile

aperta

scalabile

# il Web (1989)

## 3 principi

URL

HTTP

HTML

## Separazione tra

dati e

documenti

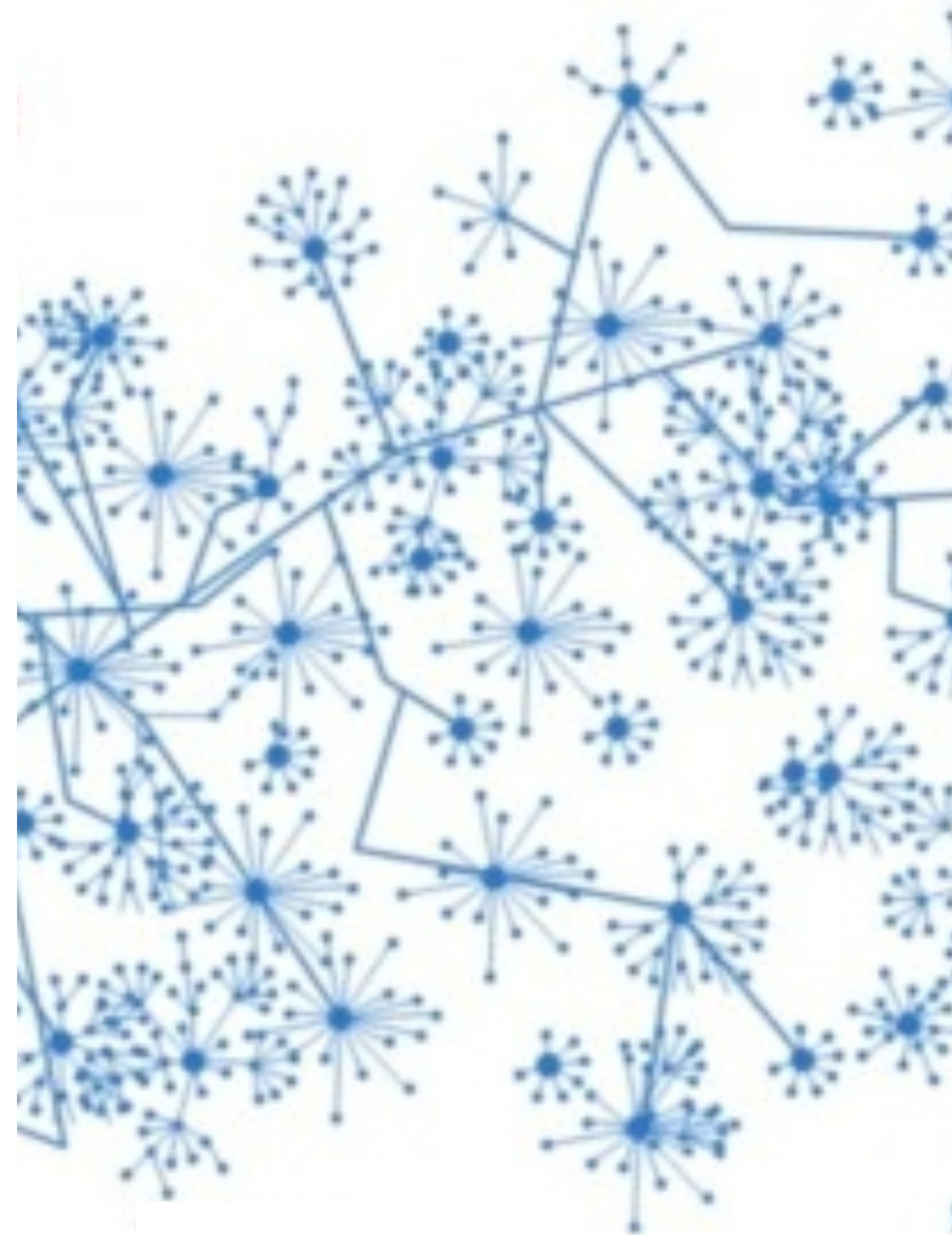
## Indipendenza da

hardware

software

scala

Adozione di **standard** per:  
la codifica dei caratteri (i18n)  
accessibilità



# Architettura del Web

Il Web è un ipertesto globale basato su 3 principi e tecnologie:

1. **URI** : Universal Resource Identifier (es. un URL)

2. **HTTP**: HyperText Transfer Protocol (on TCP/IP)



3. **HTML**: Hypertext Markup Language

T. Berners-Lee, "The World Wide Web - Past, present, future. Exploring Universality", 2002.



# HTML: esempio

```
<html>
```

```
<head>
```

```
<title>document title</title>
```

```
</head>
```

```
<body>
```

The text of the document

```
</body>
```

```
</html>
```



# da *Enquire* al Web

1989-90: Tim Berners-Lee al CERN

“When I first began tinkering with a software program that eventually gave rise to the idea of the World Wide Web, I named it *Enquire*, short of *Enquire Within upon Everything*, a musty old book of Victorian advice I noticed as a child in my parents’ house outside London. With its title suggestive of magic, the book served as *a portal to a world of information*, everything from how to remove clothing stains to tips on investing money. Not a perfect analogy for the Web, but a primitive starting point”.

“What that first bit of Enquire code led me to was something much larger, a vision encompassing the decentralized, organic growth of ideas, technology, and society.

The vision I have for the Web is about anything being potentially connected with anything. It is a vision that provides us with new freedom, and allows us to grow faster than we ever could when we were fettered by the hierarchical classification systems into which we bound ourselves. It leaves the entirety of our previous fears for the future as one set among many. And it brings the workings of society closer to the workings of our minds.”

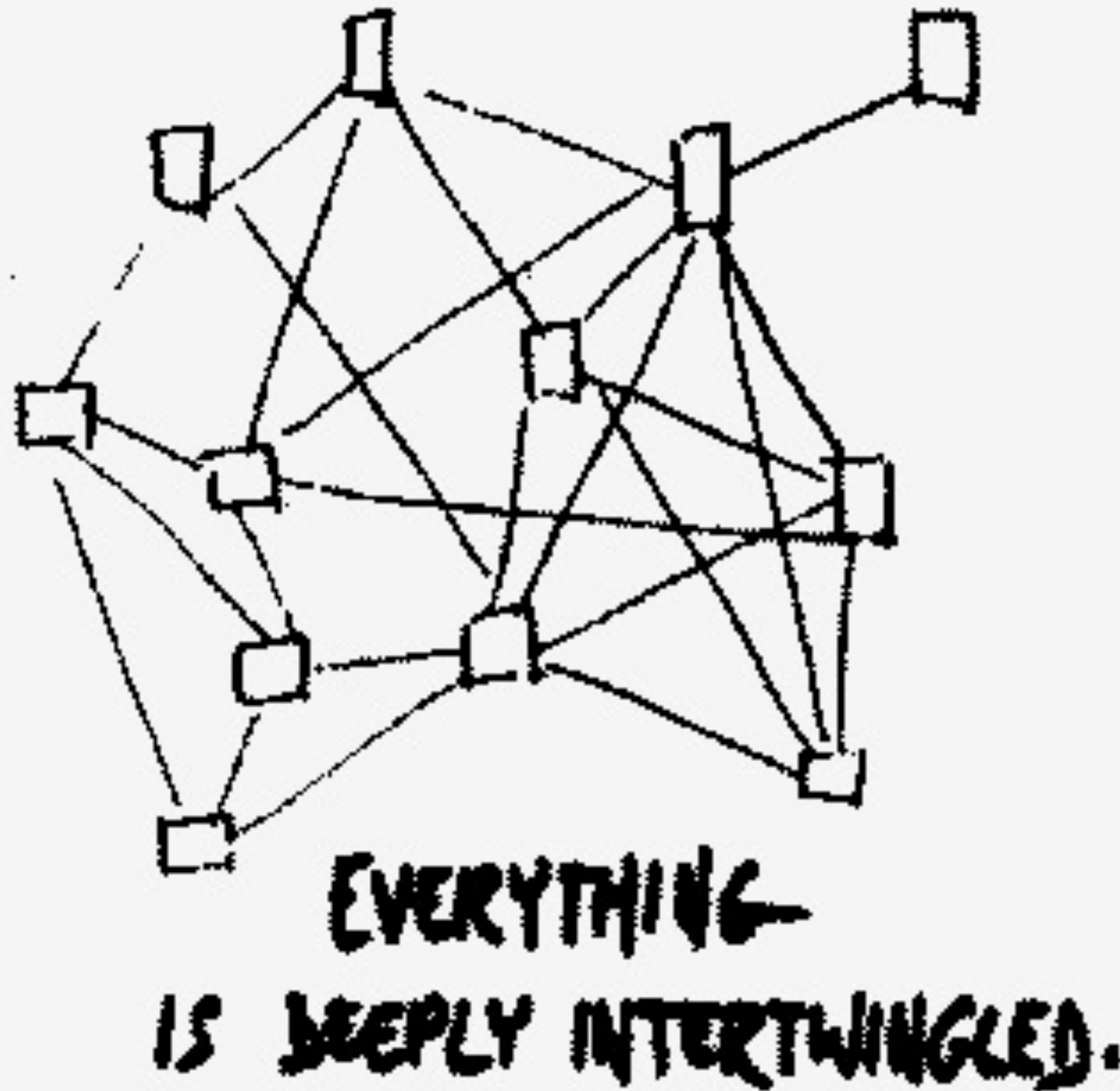
T. Berners Lee, Weaving the web, pp. 1-2.



# History



T. Nelson, Literary Machines 90.1 (1981-90)  
Inventa il concetto di ipertesto



# Un sistema di informazione universale

*“Suppose all the information stored on computers everywhere were linked, I thought. Suppose I could program my computer to create a space in which anything could be linked to anything. All the bits of information in every computer at CERN, and on the planet, would be available to me and to anyone else. There would be a single, global information space.”*

T. Berners Lee, Weaving the web, p. 4.

# Filosofia

Quello che conta sono i collegamenti liberi tra idee:

“A computer typically keeps information in rigid hierarchies and matrices, whereas the human mind has the special ability to link random bits of data. When I smell coffee, strong and stale, I may find myself again in a small room over a corner coffeehouse in Oxford. *My brain makes a link, and instantly transports me there.*”

Funziona come il sistema della citazioni scientifiche





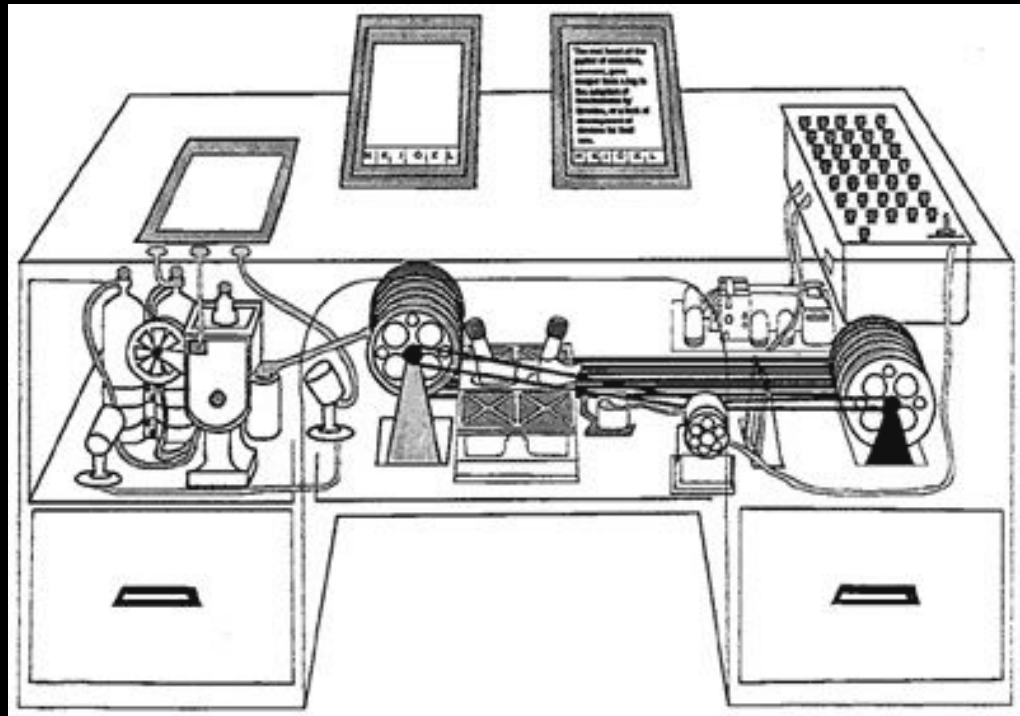
# Storia

## V. Bush, As we may think (1945)

<http://www.theatlantic.com/doc/194507/bush>

Professionally our methods of transmitting and reviewing the results of research are generations old and by now are totally inadequate for their purpose. If the aggregate time spent in writing scholarly works and in reading them could be evaluated, the ratio between these amounts of time might well be startling. ...

Mendel's concept of the laws of genetics was lost to the world for a generation because his publication did not reach the few who were capable of grasping and extending it; and this sort of catastrophe is undoubtedly being repeated all about us, as truly significant attainments become lost in the mass of the inconsequential.



il memex

The difficulty seems to be, not so much that we publish unduly in view of the extent and variety of present day interests, but rather that publication has been extended far beyond our present ability to make real use of the record.

....

But there are signs of a change as new and powerful instrumentalities come into use.



# Una nuova professione

“There is a new profession of **trail blazers**, those who find delight in the task of establishing useful trails through the enormous mass of the common record. The inheritance from the master becomes, not only his additions to the world's record, but for his disciples the entire scaffolding by which they were erected”.



# una visione più ampia

V. Bush, *Science: The Endless Frontier. A Report to the President*, Washington, U.S. Government Printing Office, 1945 [report]

## Science The Endless Frontier

**A Report to the President by Vannevar Bush,  
Director of the Office of Scientific Research and  
Development, July 1945**

(United States Government Printing Office, Washington: 1945)

### TABLE OF CONTENTS

- [Letter of Transmittal](#)
- [President Roosevelt's Letter](#)
- [Summary of the Report](#)
- [1. Introduction:](#)
  - [Scientific Progress is Essential](#)
  - [Science is a Proper Concern of Government](#)
  - [Government Relations to Science - Past and Future](#)
  - [Freedom of Inquiry Must be Preserved](#)
- [2. The War Against Disease:](#)
  - [In War](#)
  - [In Peace](#)
  - [Unsolved Problems](#)
  - [Broad and Basic Studies Needed](#)
  - [Coordinated Attack on Special Problems](#)
  - [Action is Necessary](#)
- [3. Science and the Public Welfare:](#)
  - [Relation to National Security](#)
  - [Science and Jobs](#)
  - [The Importance of Basic Research](#)
  - [Centers of Basic Research](#)
  - [Research Within the Government](#)
  - [Industrial Research](#)
  - [International Exchange of Scientific Information](#)
- [The Special Need for Federal Support](#)
- [The Cost of a Program](#)
- [4. Renewal of our Scientific Talent:](#)
  - [Nature of the Problem](#)
  - [A Note of Warning](#)
  - [The Wartime Deficit](#)
  - [Improve the Quality](#)
  - [Remove the Barriers](#)
  - [The Generation in Uniform Must Not be Lost](#)
  - [A Program](#)
- [5. A Problem of Scientific Reconversion:](#)
  - [Effects of Mobilization of Science for War](#)
  - [Security Restrictions Should be Lifted Promptly](#)
  - [Need for Coordination](#)
  - [A Board to Control Release](#)
  - [Publication Should be Encouraged](#)
- [6. The Means to the End:](#)
  - [New Responsibilities for Government](#)
  - [The Mechanism](#)
  - [Five Fundamentals](#)
  - [Military Research](#)
  - [National Research Foundation](#)
    - I. Purposes
    - II. Members
    - III. Organizations
    - IV. Functions
    - V. Patent Policy
    - VI. Special Authority
    - VII. Budget
  - Action by Congress

# 1. Guida all'uso dei motori di ricerca

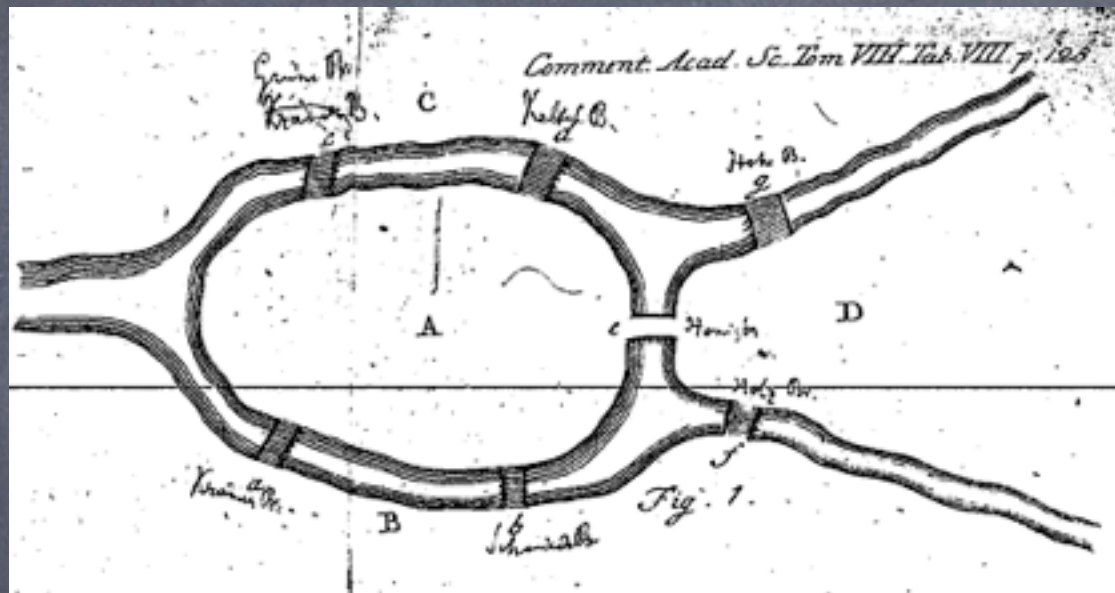
# Il Web è il nostro spazio di ricerca

Quindi, se vogliamo trovare qualcosa al suo interno, dobbiamo sapere come è fatto



# Topologia del Web

Il web è un grafo, un insieme di vertici (nodi) uniti da spigoli (link).



Grafo: i ponti di Koenigsberg

Un'importante proprietà dei grafi: l'esistenza di un percorso da un nodo a un altro non dipende dalla nostra capacità di trovarlo ma è una proprietà del grafo. (Eulero, 1736)

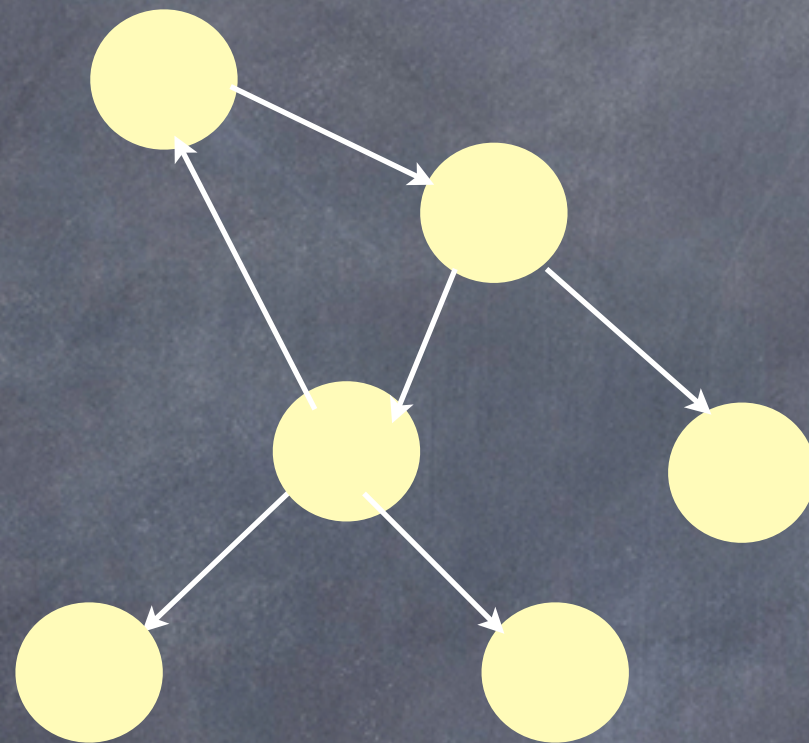


Il Web è una rete diretta,  
complessa e a invarianza di  
scala

Cioè?

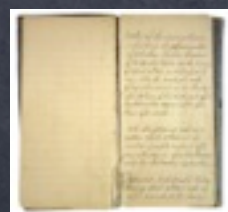


a) il web è un **grafo orientato**  
(rete diretta)

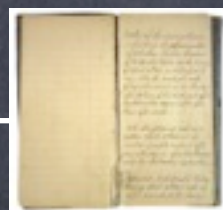


altri esempi?

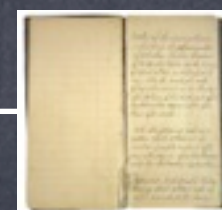
la rete della comunità scientifica  
(scritti e citazioni sono rispettivamente nodi e link)



Platone



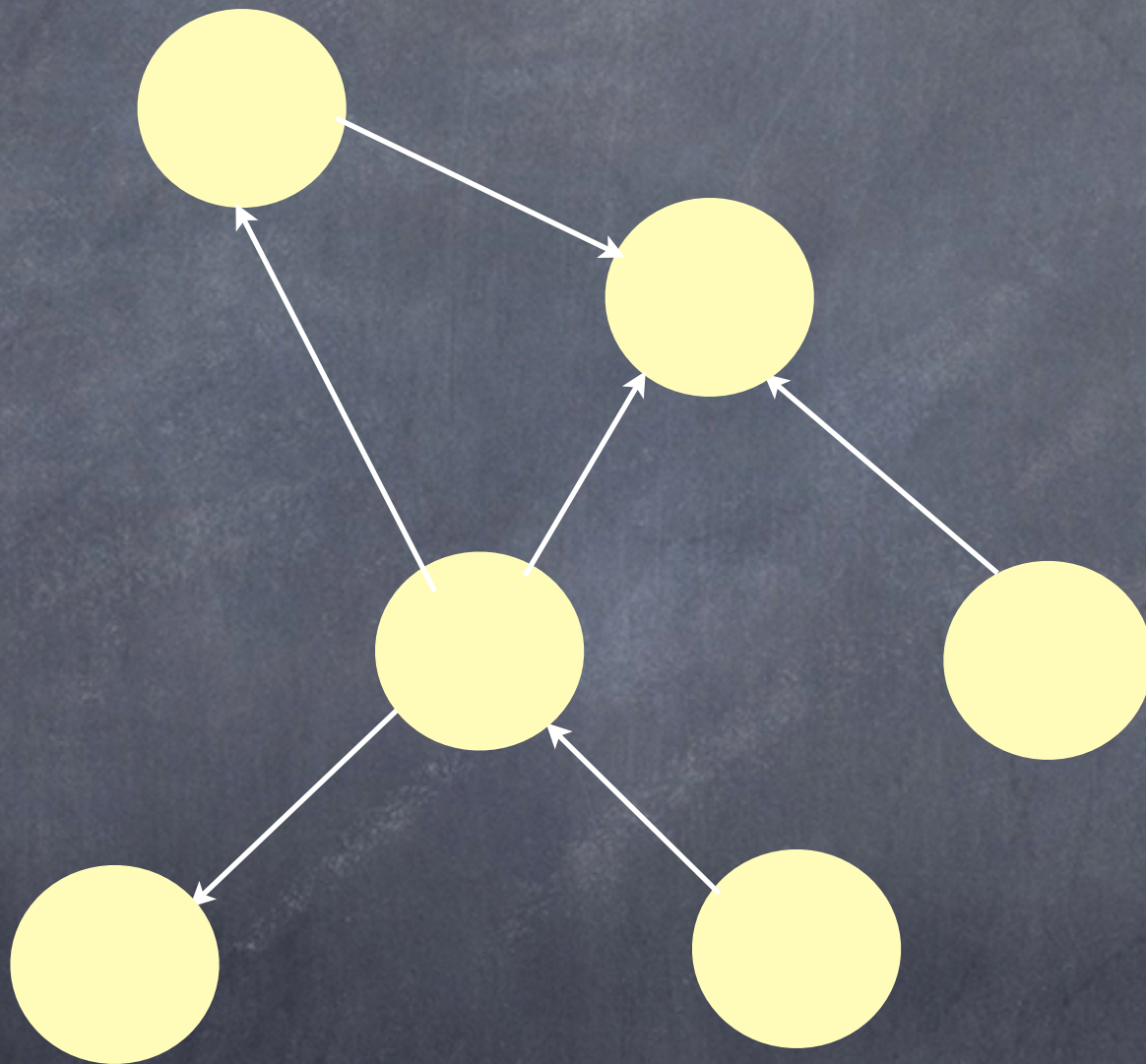
Kant



Di Donato



b) per far parte del grafo, **basta un link** per ogni nodo





## c) il web è un mondo piccolo

“Sei gradi di separazione” (Milgram): quanti contatti (link) sono necessari per connettere tra loro due individui (nodi) scelti a caso nella società?

**Risposta:** in media il numero minimo di intermediari necessari è 5,5. La società è una rete di 6 miliardi di nodi in cui la distanza media tra un nodo e l'altro non supera i 6 link.

**Un mondo piccolo è una ragnatela molto fitta** anche la rete delle pubblicazioni - citazioni scientifiche è un mondo piccolo.

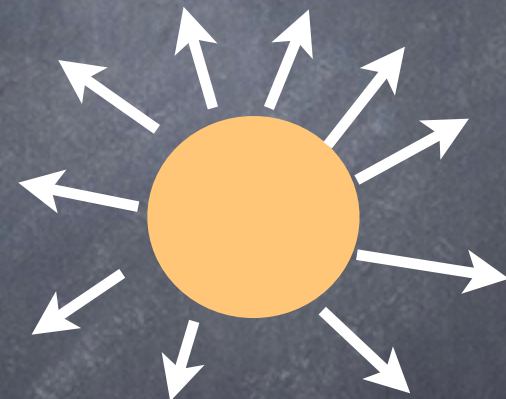


## d) presenza di connettori (hub) e authorities

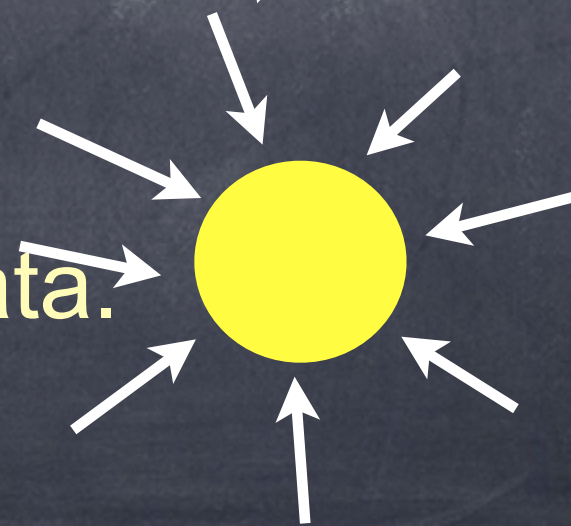
Disseminata in varie occupazioni, esiste una manciata di persone che possiede l'abilità davvero straordinaria di stringere un numero eccezionale di amicizie e conoscenze. Sono i connettori.

I connettori sono presenti in tutte le grandi reti complesse.

I connettori hanno molti link in uscita.



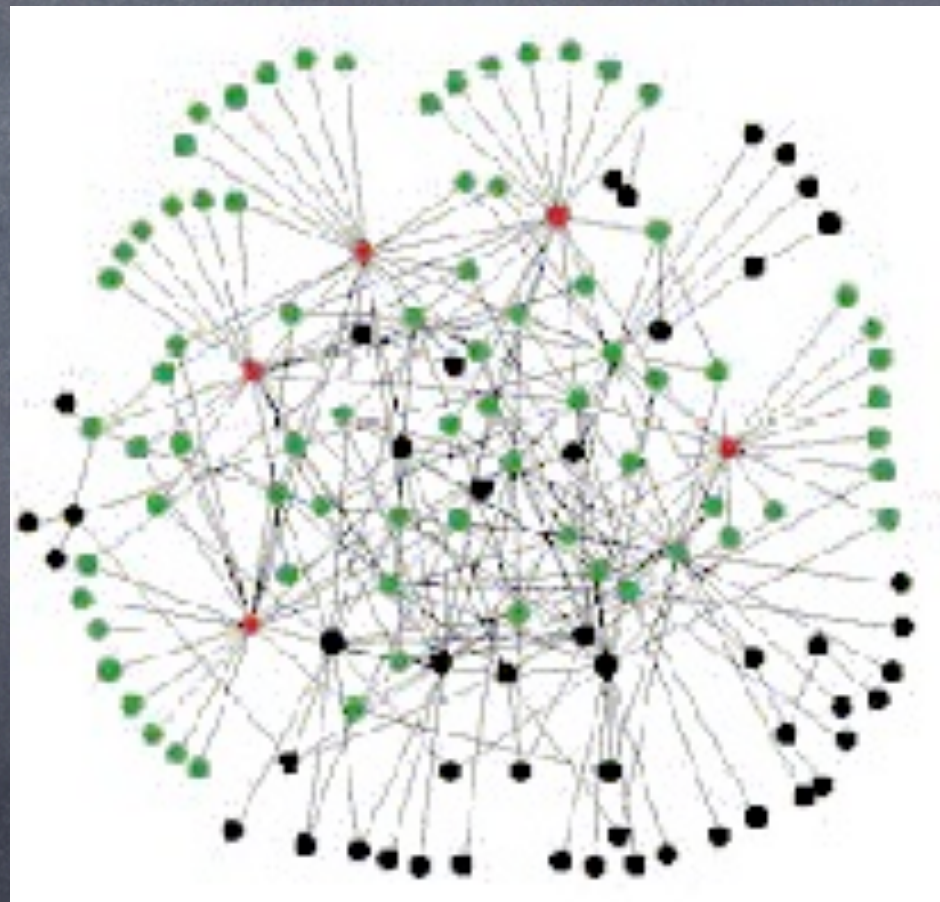
Le autorità hanno molti link in entrata.





## e) il **numero di link** misura la visibilità di un nodo

L'architettura del WWW è dominata da pochissimi nodi altamente connessi (hub) che tengono insieme molti nodi poco connessi.



A scale-free network of 130 nodes generated by the scale-free model.

The five biggest nodes are shown in red, and they are in contact with 60% of nodes (green).

<http://www.nd.edu/~networks/Linked/>

NB. La presenza di hub rende queste reti forti agli attacchi (corollario: la rete si mette in crisi attaccando gli hub)



## f) il principio 80/20

l'80% dei piselli viene dal 20% di baccelli;

l'80% dei profitti è prodotto dal 20% di impiegati;

l'80% di link parte dal 20% di pagine;

l'80% citazioni proviene dal 35% degli scritti;

etc.



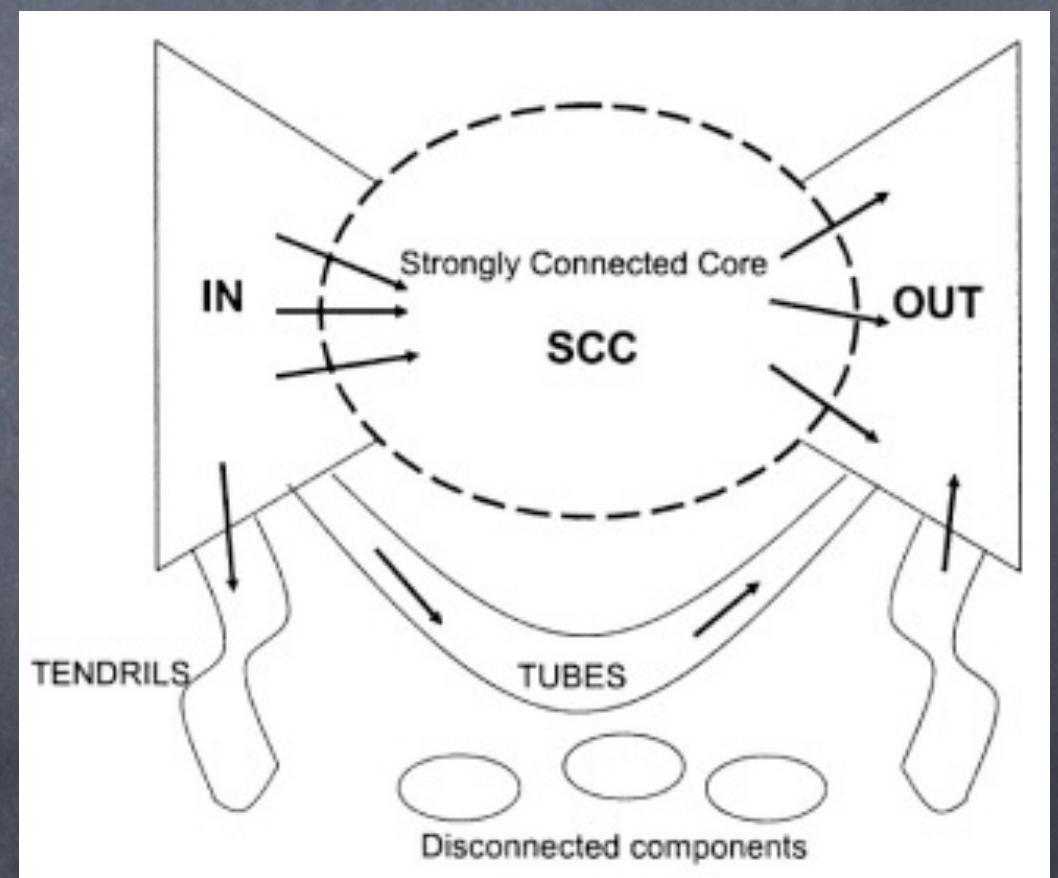
Il WWW è anche un mondo grande, perché è fatto di “**sottomondi**”. Ciò dipende dal fatto che è un grafo orientato (rete diretta). Una conseguenza di questa proprietà è che **non forma una singola rete omogenea, ma è diviso in 4 continenti di dimensioni pressoché equivalenti.**

Il primo è il **corpo centrale (SCC)**. E' facile da navigare perché molto interconnesso, esiste sempre un percorso per muoversi al suo interno.

Il secondo, detto continente **IN** è più difficile da navigare perché ha collegamenti verso SCC ma non da SCC.

Il terzo continente (**OUT**) è raggiungibile da nodi del corpo centrale, ma non ci sono link per tornare indietro (è fatto di siti “a fondo chiuso”).

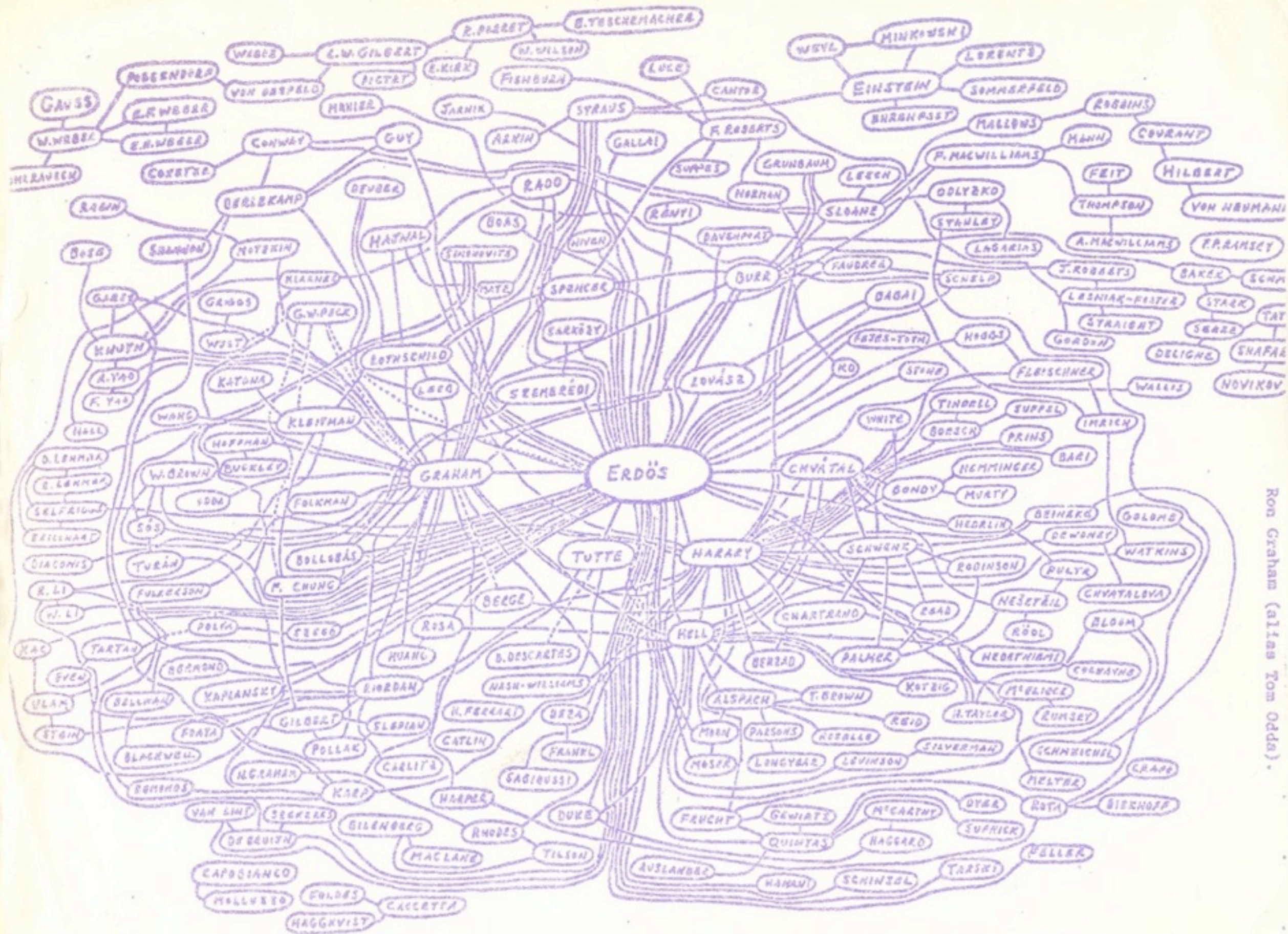
il quarto continente è fatto di **tentacoli**, **isole** separate e **tubi**, gruppi di pagine collegate tra loro ma non al corpo centrale.





NB. Anche la rete delle citazioni scientifiche è una rete diretta e obbedisce alle stesse regole





Ron Graham (alias Tom Odde).



Figure 1  
To appear in Topics in Graph Theory (P. Harary, ed.), New York Academy of Sciences (1979).



I motori di ricerca mappano solo il 50% del web, cioè il corpo centrale e il continente OUT....il resto non è navigabile.

Essere online dunque non significa, di per sé, essere visibili.

In che modo i motori trovano le pagine sul Web?





# Il più famoso

Google è stato inventato da due giovani ricercatori di teoria dei grafi, Sergej Brin e Larry Page, mentre analizzavano le caratteristiche dei grafi orientati.

Brin e Page concentrano l'analisi su un problema particolare: come tornare indietro in un grafo orientato, cioè come scoprire se un sito è stato linkato, e da chi.

Brin e Page inventano **PageRank**, l'algoritmo che consente a Google di trovare le risorse sfruttando questo limite architettonico del Web e trasformandolo in una ricchezza.



# Google un indicatore di popolarità

PageRank affronta lo stesso problema fronteggiato da Eugene Garfield pochi decenni prima\*.

PageRank un sistema per valutare la popolarità di un sito, incrociando criteri quantitativi e criteri di autorevolezza, cioè pesando i link sulla base della loro importanza. In breve, Google identifica la posizione di un nodo nel grafo e privilegia, nelle risposte alle query, le grandi autorità e i grandi connettori.

Ad esempio, è intuitivo che se la mia pagina personale viene linkata da quella di un sito di un mio collega, secondo Google l'importanza di quel collegamento è minore rispetto a quella di un link dal sito di [repubblica.it](http://repubblica.it), anche se il ricercatore che mi cita è uno scienziato illustre. Dunque ci sono link che pesano più di altri in termini di popolarità.

# esplorando la superficie del Web: sull'uso dei SES

Anche se i motori di ricerca liberamente e pubblicamente disponibili sono centinaia, pochissimi catturano la stragrande maggioranza del pubblico. Secondo la ben nota regola 80/20, **l'80 per cento degli utenti usa il 20 per cento di applicazioni.**

La percentuale è ancora più estrema con i motori di ricerca.

# “In SES we trust”

Alcune indagini hanno rivelato che più di due terzi degli utenti credono che i motori di ricerca siano una fonte d'informazione equa e imparziale.

Inoltre, gli utenti più sicuri sono quelli che hanno meno conoscenze in materia.



Un uso intelligente  
dei SES è  
essenziale per un  
buon ricercatore

# Regola n. 1

I motori sono  
tanti e diversi

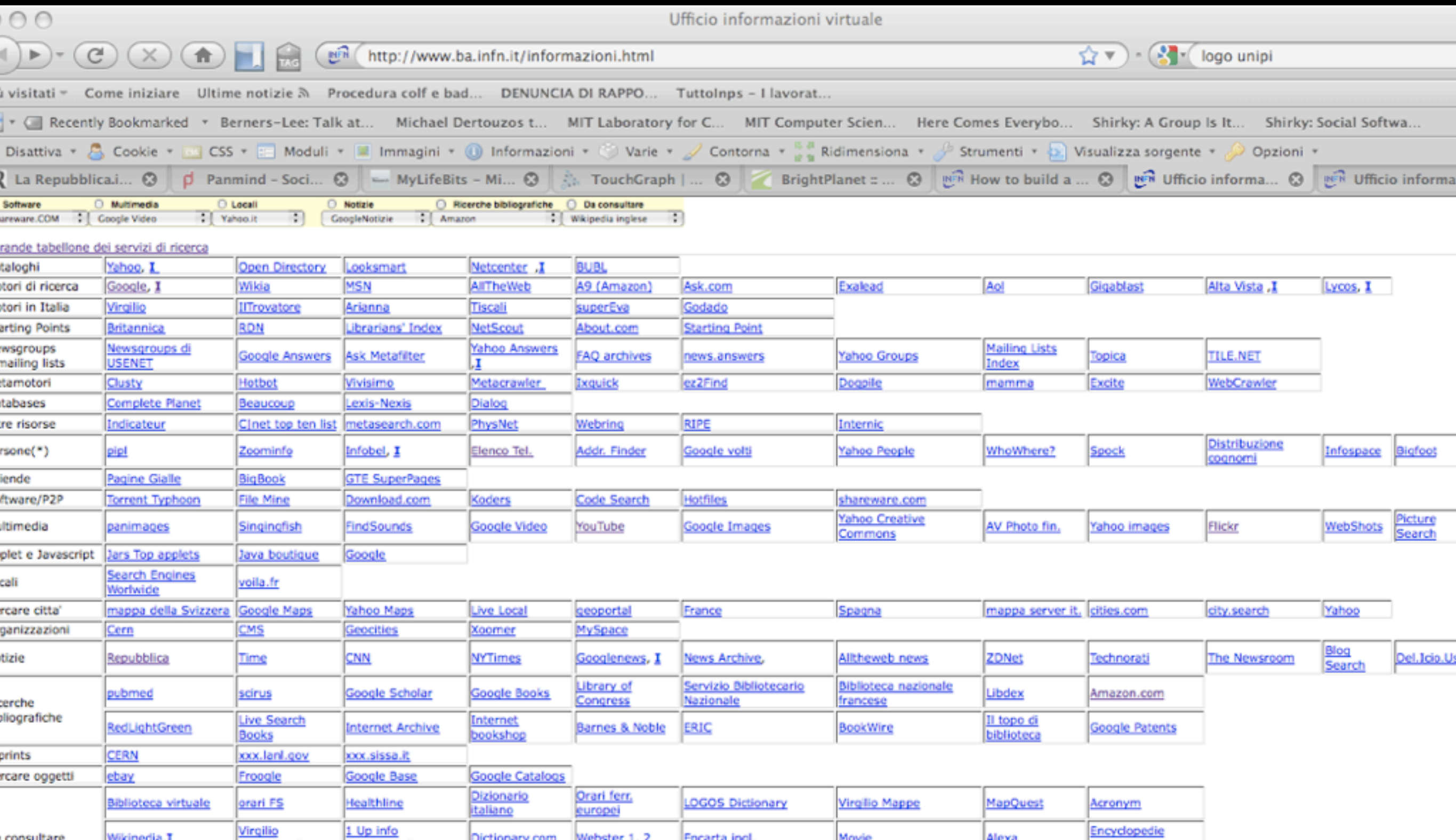


# tanti!!!

Fravia's searching <b>MAPA</b> ( <u>m</u> asks and <u>p</u> ages) <span>● : has <a href="#">cache</a></span> <span>● : special page</span> <span>● : has <a href="#">NEAR</a></span> <span>● : crippled</span> <a href="#">ss.ee.</a>						
<a href="#">operators</a>						
<a href="#">Best</a> s.e.	<a href="#">CUIL</a>	<a href="#">MSNsearch</a> ●	<a href="#">Google</a> ● <a href="#">GOOGLE</a> ●	<a href="#">Ask</a> ● (date)	<a href="#">Yahoo!</a> ● ● <a href="#">YAHOO</a> ● <a href="#">Fast</a> ● <a href="#">Altavista</a> ●	
Second Tier	<a href="#">Alexa</a> ●	<a href="#">Exalead</a> (date & regexp) ●	<a href="#">A 9</a> (google's ●)	<a href="#">Baidu</a> ●		
Useful s.e.	<a href="#">Wayback</a> (past)	<a href="#">Lycos</a>	● <a href="#">Gigablast</a> ●	<a href="#">Swicki</a> ("vertical")	<a href="#">IceRocket</a> (webarchive)	<a href="#">Rollyo</a> ("vertical")
Graph s.e.	<a href="#">Kart00</a> (graph)	<a href="#">Touch</a> (graph)	<a href="#">Ujiko</a> (graph)	<a href="#">Dicy</a> (cluster)	<a href="#">Mooter</a> (cluster)	
"Visual" s.e.	<a href="#">lyGO</a> ("visual" search)	<a href="#">yaouba</a> ("visual" & anon)	<a href="#">searchme</a> ("visual" search)			
Other	<a href="#">Entireweb</a>	<a href="#">Excite</a> (not + but very ill)	<a href="#">Factbite</a> (encycl)	<a href="#">dmoz</a> (directory)	<a href="#">Furl</a> (webarchive)	<a href="#">[FTPSEARCH]</a>
Cadavers	<a href="#">Teoma</a> (+ 07)	<a href="#">Wisenut</a> (+ 07)	<a href="#">Ouverture</a> (+ 07)	<a href="#">Northernlight</a> (+ 02)	<a href="#">Webtop</a> (+ 01)	
searching techniques and tips @ <a href="#">fravia's</a>	<a href="#">Golden rules</a>	<a href="#">Long term</a>	<a href="#">Short term</a>	<a href="#">Deep web</a>	<a href="#">Files repos</a>	
	<a href="#">Targets</a>	<a href="#">Local</a>	<a href="#">Regional</a>	<a href="#">Compound</a>	<a href="#">Usenet</a>	<a href="#">Accmail</a>
	<a href="#">Live searches</a>	<a href="#">Combing</a>	<a href="#">Klebing</a>	<a href="#">Guessing</a>	<a href="#">Databases</a>	<a href="#">Allinones</a>
	<a href="#">Images</a>	<a href="#">Books</a>	<a href="#">Laws</a>	<a href="#">Files</a>	<a href="#">Filez</a>	<a href="#">Passwords</a>

<http://www.searchlores.org/main.htm>

e ci sono molti siti e molti libri in  
materia...



I motori di ricerca usano algoritmi diversi, per cui ci restituiscono diversi risultati di una query.

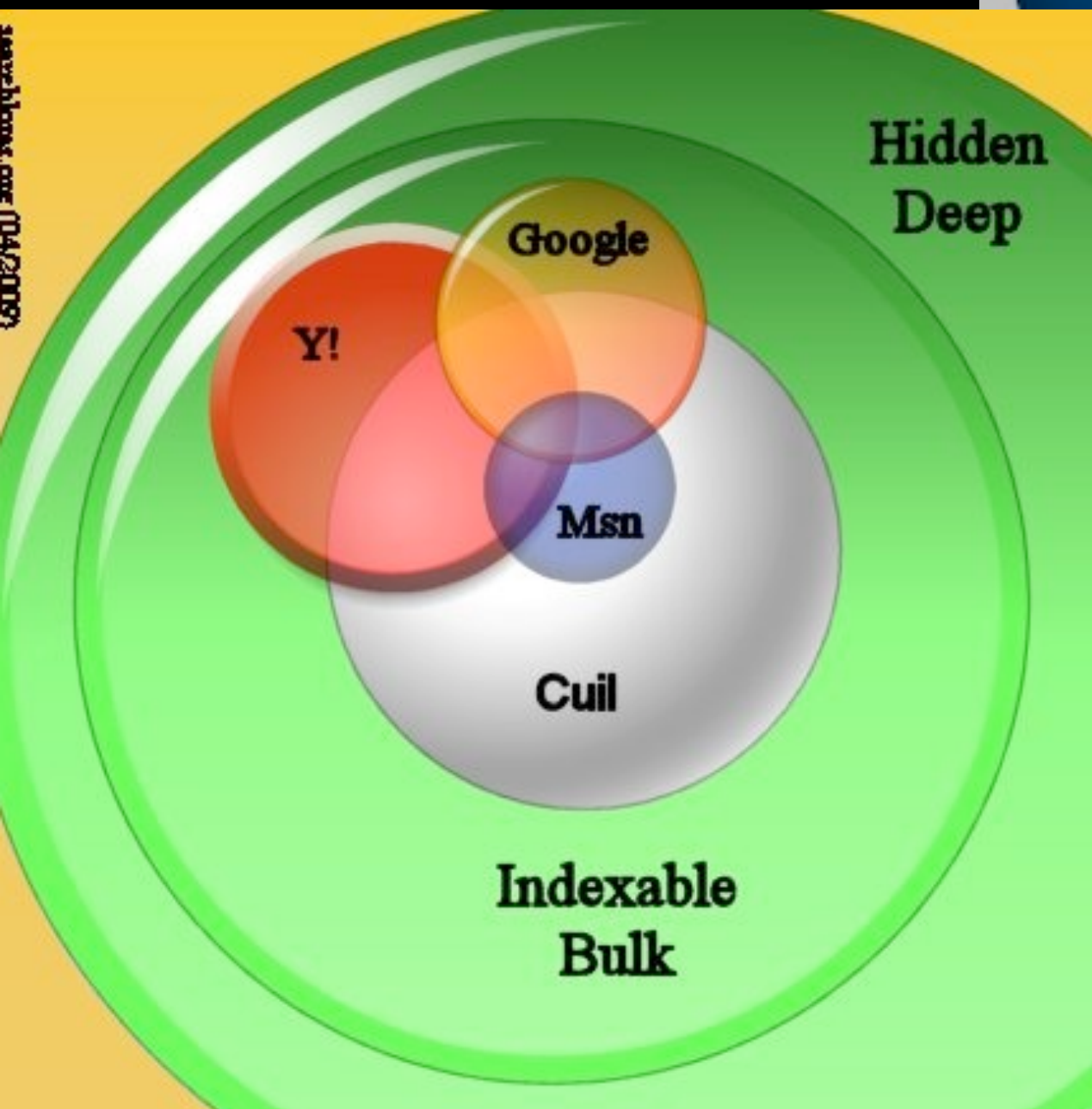
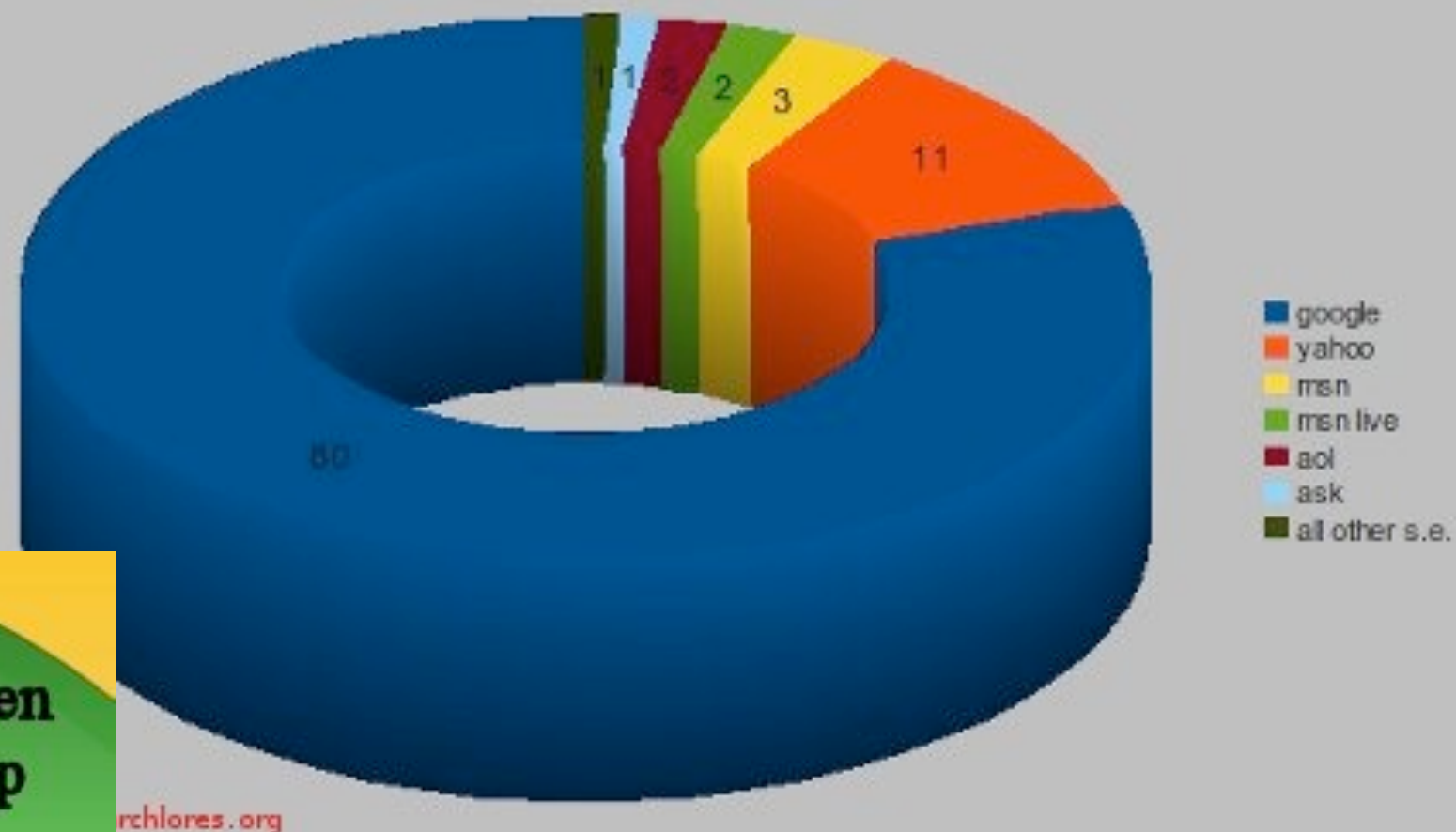
In altre parole: se vi limitate ad usare sempre e sologoogle, coprirete meno della metà del web visibile (e probabilmente neppure  $1/5 \sim 1/10$  di quello nascosto). E questo è vero nonostante il fatto che google è un buon motore di ricerca.

<http://www.searchlores.org/main.htm>



# L'uso..

Main search engines usage in March 2009 (in %)



...e la  
copertura dei  
SES

<http://www.searchlores.org/main.htm>

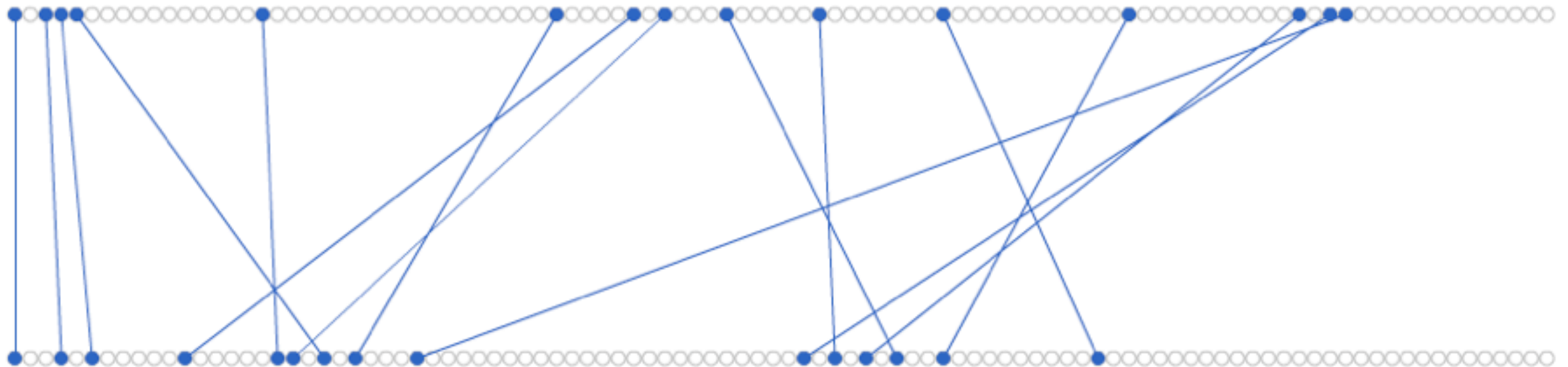
searchlores

compare 'em

Also: [Compare Google and Google.cn.](#)

Google™

touch the dots!



YAHOO!

Regola 2.

Imparare a  
formulare le  
query



# Un consiglio preliminare: Come trasformare le domande in query efficaci

## Joe A. User's 7 questions

I collected the following queries on the morning of 5 february 2003, taking them at random from [Jeeves' peek](#), where you can see what people are searching for on jeeves *eo ipso* (many similar services are listed on the [ad hoc section](#) of my sites).

Here we go with 7 questions (by "Joe A. User", where the A stays for 'average') that will allow me to show you, to-day, some searching approaches --and caveats-- you may not be aware of.

Joe A. User's 7 questions		
<a href="#">Where can I find the flag of the country Ireland?</a>	ubi	the real meaning of a query
<a href="#">Why is the sun yellow?</a>	cur	knowing how to rephrase
<a href="#">Where can I find information about the chemical element chlorine?</a>	quibus auxiliis	keeping the seeker's balance
<a href="#">How do I find the Italian national hymn?</a>	quomodo	regional searching
<a href="#">When was the battle of Kadesh fought?</a>	quando	evaluation lore (1)
<a href="#">Who is the father of modern biotechnology?</a>	quis	evaluation lore (2)
<a href="#">What is the neurological condition acoustic neuroma?</a>	quid	local and specialized engines

As you can see, all these questions have been presented in normal, plain English. Now let's have a look at the answers given by "ask jeeves", at the answers you should be able to find quickly, and at the possible approaches to find SIMILAR answers on the web. Let us see also whether the algorithms used by the various search engines are good enough as they are, are miserable, could be improved, should be ditched, or whatever.

# Regola 3.

## Usare gli operatori

<http://www.searchlores.org/operators.htm>

Ex: Google operators

site:

allintitle: (all of the query words in the title)

intitle: (that word in the title)

allinurl: (all of the query words in the URL)

**inurl: (that word in the URL)**

cache:

link:

related: (pages that are "similar" to a specified web page)

info: (google's info)

# altri consigli pratici

1. usate le minuscole
2. usate le virgolette [“”]
3. inserite errori
4. usate i booleani [ AND | & | + ] and [ AND NOT ||| - ]
5. usate l'asterisco [\*]



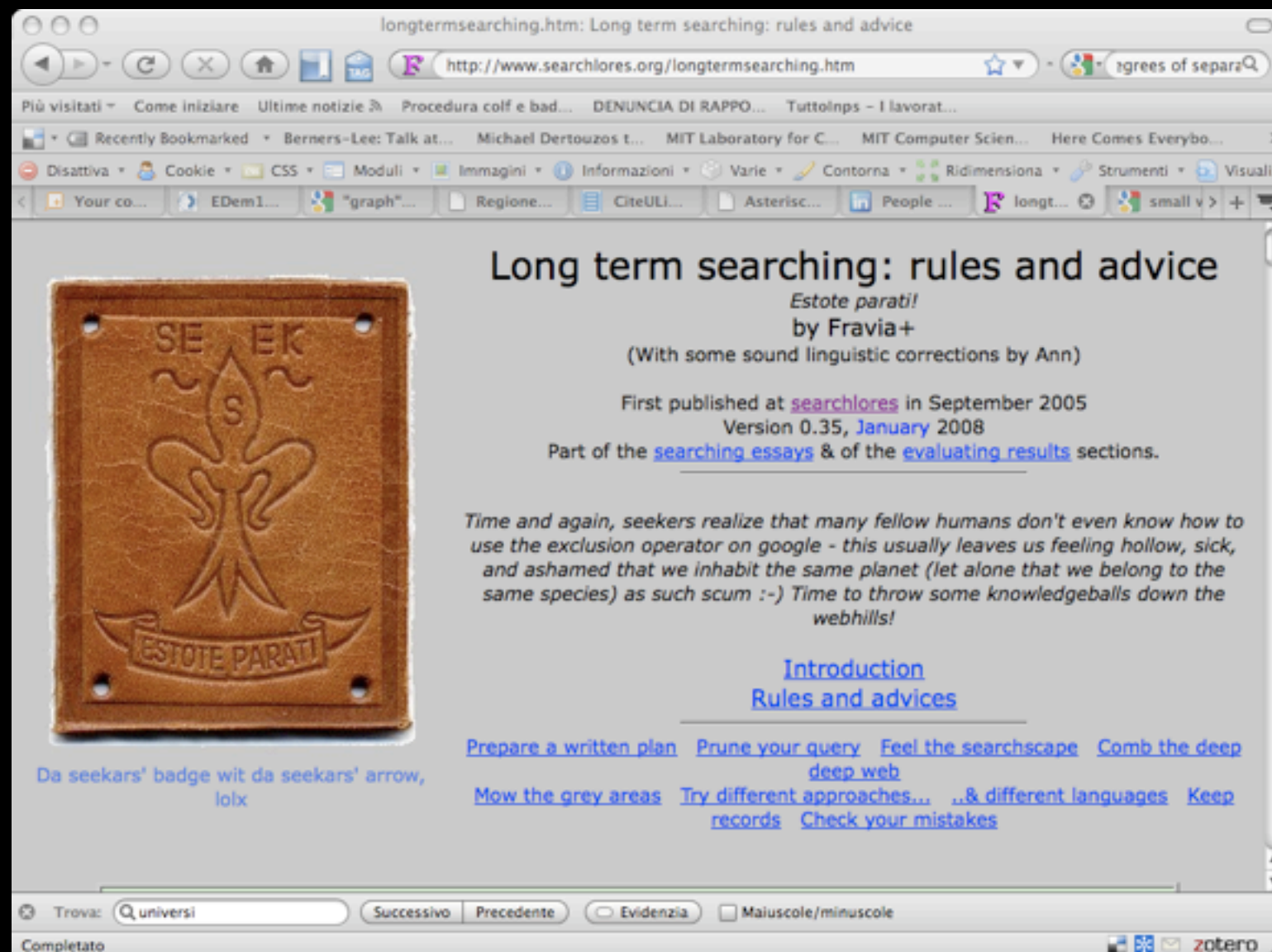
# Consigli/link utili

<http://www.smashingapps.com/2008/07/18/13-most-unusual-search-engines-you-should-remember.html>

<http://www.seochat.com/seo-tools/graphical-search-engine-comparison/>

# Ricerche di lunga durata

<<http://www.searchlores.org/longtermsearching.htm>>



see also: [http://www.searchlores.org/effective\\_searching.htm](http://www.searchlores.org/effective_searching.htm)

1. preparate un piano (scritto)

brainstorming; parole chiave

2. semplificate la vostra query!

es.

"title:index title:of" -originurlextension:htm

-originurlextension:html -papers -copyright +Oreilly

usate almeno 3 motori di ricerca diversi

3. fate alcune ricerche preliminari per  
verificare di aver fatto le scelte giuste



# Una regola importante:

**Fate un cambiamento alla volta!**

Inserendo più modifiche assieme, non sarete in grado di capire quale di essere è stata determinante (e sarà anche più difficile tornare indietro).

# Il deep Web

Con l'espressione deep web ci si riferisce a quei contenuti che non fanno parte della sua superficie, che non è indicizzata dai motori di ricerca.

**Il deep Web è diversi ordini di grandezza più grande della sua superficie** (a settembre 2001 la differenza veniva stimata essere tra le 400 e le 550 volte).

**Le risorse del deep web** rientrano in una delle seguenti categorie:

- \* Contenuti dinamici
- \* Contenuti che non sono linkati da nessun'altra pagina
- \* Web privato o ad accesso limitato
- \* Web contestuale
- \* Contenuti “scripted”
- \* Contenuti non-HTML



**Le digital library** sono un esempio di deep web.

Qualche esempio:

- <http://burckhardtsource.org>
- <http://wittgensteinsource.org>
- <http://gramscisource.org>

4. esplorate il deep (deep) web

5. identificate le "aree grigie"

Le aree grigie del web:

- conference papers and proceedings,
- unpublished dissertations on relevant topics,
- "unofficial" messageboards,
- IRC channels,
- blogs

6. provate approcci diversi

7. rifate le vostre query in lingue diverse

<http://www.searchlores.org/regional.htm>

<http://www.searchlores.org/dictiona.htm>

<<http://www.searchlores.org/longtermsearching.htm>>

## 8. tenete traccia delle vostre attività

A questo scopo vi saranno molto utili strumenti che vedremo più tardi, come

- **delicious**
- **citeulike**

Infine, rivedete le vostre query (in cerca di errori)



## Misspellings of *Britney Spears* Typed into Google

			488941	britney spears			
40134	brittany spears	544	brittaney spears	121	britainy spears	73	pritney spears
36315	brittney spears	544	brittnay spears	121	britmey spears	66	brintany spears
24342	britany spears	364	britey spears	109	brietney spears	66	britnery spears
7331	britny spears	364	brittiny spears	109	brithny spears	59	briitney spears
6633	briteny spears	329	brtney spears	109	britni spears	59	britinay spears
2696	britteny spears	269	bretney spears	109	brittant spears	54	britneay spears
1807	briney spears	269	britneys spears	98	bittney spears	54	britner spears
1635	brittny spears	244	britne spears	98	brithey spears	54	britney's spears
1479	brintey spears	244	brytney spears	98	brittiany spears	54	britnye spears
1479	britanny spears	220	breatney spears	98	btitney spears	54	britt spears
1338	britiny spears	220	britiany spears	89	brietny spears	54	brttany spears
1211	britnet spears	199	britnney spears	89	brinety spears	48	bitany spears
1096	britiney spears	163	britnry spears	89	brintny spears	48	briny spears
991	britaney spears	147	breatny spears	89	britnie spears	48	brirney spears
991	britnay spears	147	brittiney spears	89	brittey spears	48	britant spears
811	brithney spears	147	britty spears	89	brittnet spears	48	britnety spears
811	brtiney spears	147	brotney spears	89	brity spears	48	brittanny spears
664	birtney spears	147	brutney spears	89	ritney spears	48	brttney spears
664	brintney spears	133	britteney spears	80	bretny spears	44	birttany spears
664	briteny spears	133	briyney spears	80	britnany spears	44	brittani spears
601	bitney spears	121	bittany spears	73	brinteny spears	44	brityney spears
601	brinty spears	121	bridney spears	73	brittainy spears	44	brtitney spears

altre risorse





Click logo to [enter](#) [«](#) [†](#) [»](#) [μ](#) [·](#) [•](#) [Ⓟ](#)

[www.searchlores.org](http://www.searchlores.org) (Oz & States)

[www.fravia.com](http://www.fravia.com) (Europe)

Web



## Searchlores

Advanced Internet searching [strategies](#) & [advice](#)  
[Resources](#) for [basic](#), [advanced](#) & [deep web](#) seekers

Lore [updated](#) for the [last time](#) 1 month and 6  
days ago

[Welcome](#) | [What's new](#) | [Warning](#)

[Mirrors of...](#) [Helping...](#) [Searching...](#) ...this site

You are **not** using the [Opera](#) browser... enjoy my site nevertheless.

Notice: [this site is no longer updated](#), but there is still  
a considerable amount of sheer, raw knowledge.

However, if your attention span is too short, please, by all means, [bugger off](#)

Seeking & reversing since 1995

Seekers' [Quickforms](#), [Webbits](#) & [Rose](#)

# filosofia

“... It's up to the visitor to study it's workings to take advantage of the information that is provided here.

You will also have plenty of opportunities to leave it for intriguing paths to excellent destinations. It could even happen that you wont be able to find some parts of this site that are not 'on prime real estate'. Many visitors may leave for ever. Good riddance, others will come: the web is a Quicksand.”



# struttura del sito

4 sezioni principali:

the basic pages (basic)

<<http://www.searchlores.org/basic.htm>>

the advanced pages (advanced)

<<http://www.searchlores.org/advanced.htm>>

the classroom pages (classrooms)

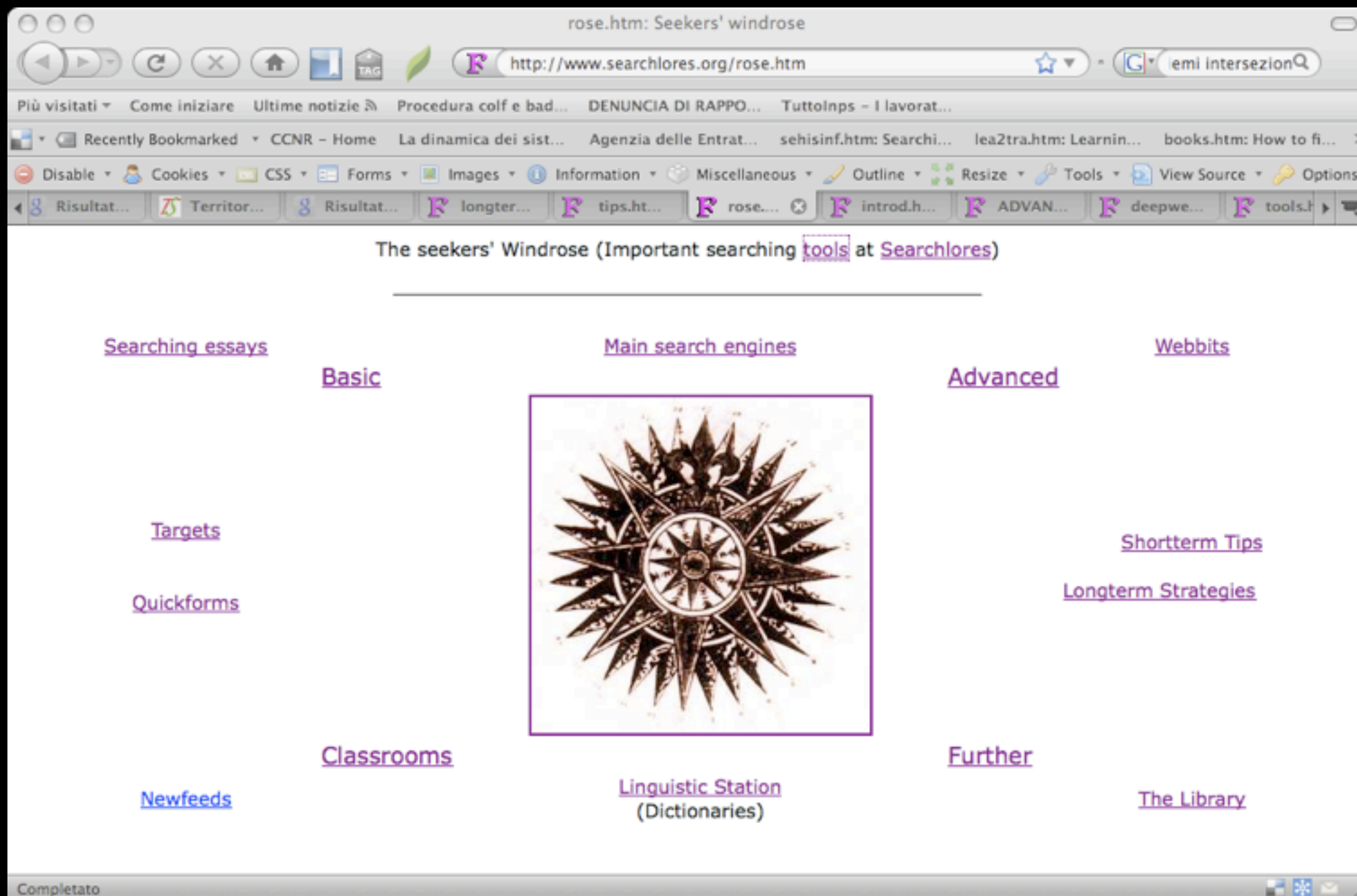
<<http://www.searchlores.org/classroo.htm>>

the farewell pages (further)

<<http://www.searchlores.org/further.htm>>

e inoltre...

# la rosa dei venti del “cercatore”





risorse on-line

Deep web searching.

The lore of searching: how to exploit the shallow deep\_web

[http://www.searchlores.org/deepweb\\_searching.htm](http://www.searchlores.org/deepweb_searching.htm)

How to find books and texts

<http://www.searchlores.org/books.htm>

Combing

<http://www.searchlores.org/combing.htm>

Regional search engines

<http://www.searchlores.org/regional.htm>

Blog

<http://www.searchlores.org/blog.htm>

Essays

<http://www.searchlores.org/essays.htm>

Classrooms

[http://www.searchlores.org/c\\_intro.htm](http://www.searchlores.org/c_intro.htm)

Conferences and workshops

<http://www.searchlores.org/mines.htm>

The lore of (dinosauria) researching (An "how to" for young web seekers)

How to research, evaluate and collate web material

by A+heist (heavily edited by fravia+), February 2008

[http://www.searchlores.org/how\\_to\\_research.htm](http://www.searchlores.org/how_to_research.htm)



# Testi per le scienze umane e sociali

<http://www.gutenberg.org>: **Project Gutenberg**

<http://gutenberg.net.au/>: Gutenberg **australia**

(As the oz law is just 50 years max...)

<http://gallica.bnf.fr/>: **Gallica**

<http://about.eserver.org/>: **Eserver**

<http://books.google.com/books?>: **Google books**

<http://scholar.google.com/>: **Google scholar**

<http://en.scientificcommons.org/>: **Index of OAI-compliant papers**

<http://www.archive.org/search.php?query=subject%3A%22search%22>: **Internet Archive**

<http://vlib.org.uk/>: **The WWW Virtual Library**

<http://digital.library.upenn.edu/books/search.html>: **The University of Pennsylvania Online Books Page**

<http://abu.cnam.fr/index.html>: **ABU: la Bibliothèque Universelle**

<http://www.opencontentalliance.org/>: **Open Content Alliance**

<http://www.readprint.com/>: Our website offers thousands of free books for students, teachers, and the classic enthusiast.

<http://www.gutenberg.org/>: There are 17,000 free books in the Project Gutenberg Online Book Catalog.

<http://www.bibliomania.com/>: Free Online Literature with more than 2000 Classic Texts

<http://digital.library.upenn.edu/books/>: Upenn.edu, Listing over 25,000 free books on the Web

<http://www.ipl.org/div/subject/browse/hum60.60.00/>: The Internet Public Library, Literature Online Texts

<http://www.literature.org/>: An Online Library of Literature. Read. Learn. Think.

<http://www.loc.gov/>: The Library of Congress serves as the research arm of the US-Congress.

<http://ota.ahds.ac.uk/>: The Oxford Text Archive hosts AHDS Literature, Languages and Linguistics.

<http://bcdlib.tc.ca/links-subjects.html>: British Columbia digital library: The focus of this set of links is on collections of electronic texts (not individual titles) preserved through libraries, archives, museums and corporate or private initiatives.

[http://un2sg4.unige.ch/athena/html/fran\\_fr.html](http://un2sg4.unige.ch/athena/html/fran_fr.html): Textes d'auteurs d'expression française

<http://www.intratext.com/>: Full-text Digital Library committed to accuracy, accessibility and usability, offering texts and corpora as lexical hypertexts

<http://www.fordham.edu/halsall/sbook2.html>: Internet Medieval Sourcebook

<http://pomoerium.com/links.htm>: classics resources

[http://un2sg4.unige.ch/athena/html/fran\\_fr.html](http://un2sg4.unige.ch/athena/html/fran_fr.html): Textes d'auteurs d'expression française

<http://www.intratext.com/>: Full-text Digital Library committed to accuracy, accessibility and usability, offering texts and corpora as lexical hypertexts

<http://www.fordham.edu/halsall/sbook2.html>: Internet Medieval Sourcebook

<http://pomoerium.com/links.htm>: classics resources

<http://germazope.uni-trier.de/Projects/DWB>: Das Deutsche Wörterbuch von Jacob und Wilhelm Grimm auf CD-ROM und im Internet

<http://www.ikp.uni-bonn.de/kant/>: Das Bonner Kant-Korpus. Elektronische Edition der Gesammelten Werke Immanuel Kants

<http://dewey.library.upenn.edu/sceti>: SCETI. Virtual facsimiles of rare books and manuscripts.

<http://dsal.uchicago.edu/index.html>: Digital South Asia Library  
<http://www.fiu.edu/~mirandas/cardinals.htm>: The Cardinals of the Holy Roman Church  
<http://www.1911encyclopedia.org/>: (Britannica, eleventh edition)  
<http://www.newadvent.org/cathen/>: The Catholic Encyclopedia  
<http://plato.stanford.edu/>: Stanford Encyclopedia of Philosophy  
[http://ourworld.compuserve.com/homepages/cornwall\\_business\\_systems/index.htm](http://ourworld.compuserve.com/homepages/cornwall_business_systems/index.htm):  
A Smaller Classical Dictionary of Biography, Mythology and Geography  
<http://lexicorient.com/e.o/index.htm>: The Encyclopaedia of the Orient.  
[http://astronomy.nju.edu.cn/twkp/astrobook/Oth\\_Historical.html](http://astronomy.nju.edu.cn/twkp/astrobook/Oth_Historical.html): Astronomical Books  
Online  
<http://www.biblegateway.com/>: "Enter the Bible passage (e.g. John 3:16),  
keyword (e.g. Jesus, prophet, etc.) or topic (e.g. salvation) you want to find"  
<http://www.questia.com/Index.jsp>: Questia, Your Online Library for Research.  
Search over 60,000 Scholarly Books and 1,000,000 Journals.

## Leggi e documenti UE e ONU

<http://www.searchlores.org/laws.htm>  
[http://www.searchlores.org/frav\\_eu2.htm](http://www.searchlores.org/frav_eu2.htm)  
<http://www.searchlores.org/eurosearch.htm>  
[http://www.searchlores.org/frav\\_eu1.htm](http://www.searchlores.org/frav_eu1.htm)

## Altre fonti

<http://avaxhome.ws/>  
<http://gigapedia.com/>

Vedi anche: <http://www.searchlores.org/books.htm>

Universal library: <http://www.searchlores.org/universallibrary.htm>



# Classicalia

<http://www.thelatinlibrary.com/index.html>: the latin library, latin

<http://www.thelatinlibrary.com/neo.html>: the latin library, neo-latin

[http://www.forumromanum.org/literature/authors\\_a.html](http://www.forumromanum.org/literature/authors_a.html): Corpus scriptorum latinorum

[http://www.fh-augsburg.de/~harsch/a\\_chron.html#latmed](http://www.fh-augsburg.de/~harsch/a_chron.html#latmed): BIBLIOTHECA AUGUSTANA  
(Bibliotheca Latina scriptorum latinorum collectio)

<http://www.corpusthomisticum.org/iopera.html>: CORPUS THOMISTICUM S. THOMAE DE  
AQUINO OPERA OMNIA

<http://www.textkit.com/>: Textkit is the Internet's largest provider of free and fully  
downloadable Greek and Latin grammars and readers. With currently 146 free books to  
choose from.

[http://www.molfettanet.com/tradizioni/pesi\\_e\\_misure.htm](http://www.molfettanet.com/tradizioni/pesi_e_misure.htm): Pesi e misure nell'antichità  
(Italian)

<http://www.ancientlibrary.com/wcd/>: The wiki classical dictionary at ancientlibrary (currently  
down :-(

<http://www.perseus.tufts.edu/hopper/>: Perseus Digital Library

# De re orientalia

<http://omphaloskepsis.com/collection/index.html>: Omphaloskepsis

Omphaloskepsis provides free access to important works of eastern literature in  
digital format.

# bibliografia

**A-L. Barabási**, *Linked: The New Science of Networks* (Perseus, Cambridge, MA, 2002).

**I.H. Witten, M. Gori, T. Numerico**, *Web Dragons. Inside the Myths of Search Engine Technology*, Morgan Kaufmann Publishers-Elsevier, San Francisco 2007.

**Fravia, Searchlores**. *Advanced Internet searching strategies & advice. Resources for basic, advanced & deep web seekers*, <http://www.searchlores.org>.

### 3. Creare, gestire, condividere bibliografie scientifiche



# Web 2.0

**“Web 2.0** is a set of economic, social, and technology trends that collectively form the basis for the next generation of the Internet - a more mature, distinctive medium characterized by user participation, openness, and network effects”.

Tim O'Reilly, 2004.

# Web 2.0: una definizione, sei concetti

## 1) Il Web come piattaforma

Es. Google account <http://www.google.it>

## 2) Software as a service (e non un prodotto)

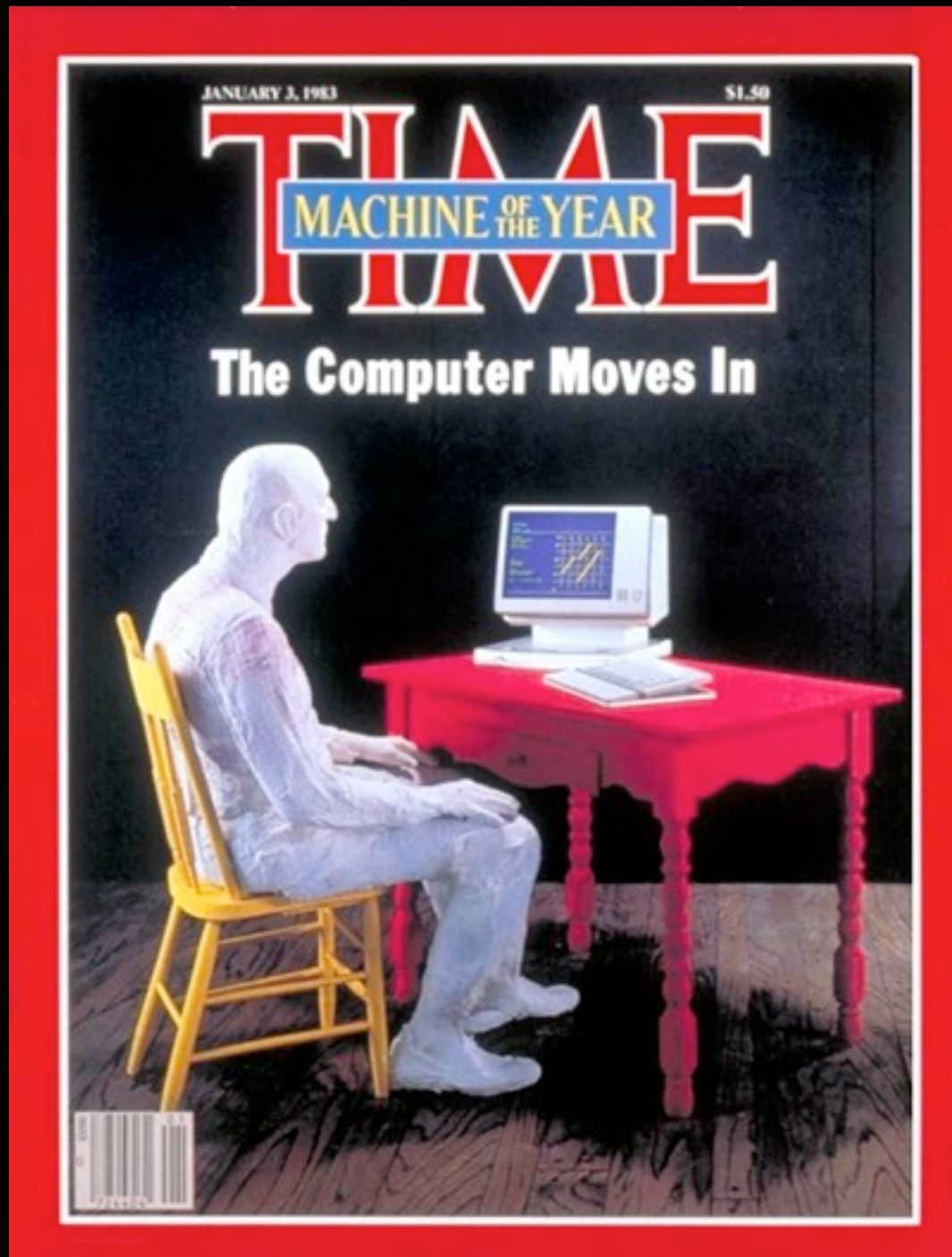
## 3) Decentralizzazione: ogni client è un server (P2P)

Es. Napster, Emule, etc..



# 4) User Generated Contents

*Gennaio 1983*



*Gennaio 2007*



# 5) intelligenza collettiva

## \* social network

# Es: Facebook, Youtube

# \* Wikipedia

## \* Blogs

## \* folksonomie

# Es. Delicious



# 6) Alcuni diritti riservati (copyleft)

## Le licenze Creative commons



The screenshot shows a web browser window with the URL <http://creativecommons.org/about/licenses/>. The page title is "Licenses - Creative Commons". The browser's address bar shows the URL, and the search bar contains the word "hypertext". The page features a green header with the Creative Commons logo and navigation links: About, News, Donate, FAQ, Wiki, Projects, Store, and International. The main content area is titled "About Licenses" and includes a search bar. Below the title, there is a paragraph explaining the six main licenses. The "License Conditions" section lists four conditions: Attribution by (person icon), Share Alike sa (circular arrow icon), Non-Commercial nc (dollar sign with slash icon), and No Derivative Works nd (equals sign icon). Each condition has a brief description. The right sidebar contains a list of links: Consulting, Who Uses CC?, What is CC?, History, Downloads, Press Room, Stay Informed, Videos and Comics, People, Alumni, Board of Directors, Fellows and Interns, Opportunities, The Licenses, Newsletter, Events, Supporters, Jobs, Developers, Case Studies, CC Talks With..., Education, Science, and Contact. The footer includes a search bar with the text "Trova: universi" and navigation buttons: Successivo, Precedente, Evidenzia, and Maiuscole/minuscole.

**About Licenses**

The following describes each of the six main licenses offered when you choose to [publish your work](#) with a Creative Commons license. We have listed them starting with the most accommodating license type you can choose and ending with the most restrictive license type you can choose.

**License Conditions**

Creators choose a set of conditions they wish to apply to their work.

- Attribution by**  
You let others copy, distribute, display, and perform your copyrighted work — and derivative works based upon it — but only if they give credit the way you request.
- Share Alike sa**  
You allow others to distribute derivative works only under a license identical to the license that governs your work.
- Non-Commercial nc**  
You let others copy, distribute, display, and perform your work — and derivative works based upon it — but for non-commercial purposes only.
- No Derivative Works nd**  
You let others copy, distribute, display, and perform only verbatim copies of your work, not derivative works based upon it.

**The Licenses**



# Scegli la licenza!

## scegli la licenza

Scegliendo una licenza Creative Commons **mantieni i tuoi diritti d'autore** ma permetti agli altri di [copiare e distribuire la tua opera](#) purché loro [te ne riconoscano la paternità](#) — e solamente sotto le condizioni che tu specifichi qui. Se vuoi offrire la tua opera senza alcuna restrizione all'uso scegli [la donazione al pubblico dominio](#).

### Allow commercial uses of your work?

☒ Si ⓘ

☐ No ⓘ

### Allow modifications of your work?

☒ Si ⓘ

☐ Yes, as long as others share alike ⓘ

☐ No ⓘ

Giurisdizione della tua licenza ⓘ

Italia ⓘ

### Additional Information

The additional fields are **optional**, but will be embedded in the HTML generated for your license. This allows users of your work to determine how to attribute it or where to go for more information about the work. ⓘ

Seleziona il formato della tua opera:

Altro ⓘ

Title of work

ⓘ

Attribute work to name

ⓘ

Attribute work to URL

ⓘ

Source work URL

ⓘ

More permissions URL

ⓘ

Scegli una licenza

## Attribution Non-commercial No Derivatives (by-nc-nd)

This license is the most restrictive of our six main licenses, allowing redistribution. This license is often called the “free advertising” license because it allows others to download your works and share them with others as long as they mention you and link back to you, but they can’t change them in any way or use them commercially.

[Read the Commons Deed](#) | [View Legal Code](#)



[Choose by-nc-nd license](#)

## Attribution Non-commercial Share Alike (by-nc-sa)

This license lets others remix, tweak, and build upon your work non-commercially, as long as they credit you and license their new creations under the identical terms. Others can download and redistribute your work just like the by-nc-nd license, but they can also translate, make remixes, and produce new stories based on your work. All new work based on yours will carry the same license, so any derivatives will also be non-commercial in nature.

[Read the Commons Deed](#) | [View Legal Code](#)



[Choose by-nc-sa license](#)

## Attribution Non-commercial (by-nc)

This license lets others remix, tweak, and build upon your work non-commercially, and although their new works must also acknowledge you and be non-commercial, they don’t have to license their derivative works on the same terms.

[Read the Commons Deed](#) | [View Legal Code](#)



[Choose by-nc license](#)

## Attribution No Derivatives (by-nd)

This license allows for redistribution, commercial and non-commercial, as long as it is passed along unchanged and in whole, with credit to you.

[Read the Commons Deed](#) | [View Legal Code](#)



[Choose by-nd license](#)

## Attribution Share Alike (by-sa)

This license lets others remix, tweak, and build upon your work even for commercial reasons, as long as they credit you and license their new creations under the identical terms. This license is often compared to open source software licenses. All new works based on yours will carry the same license, so any derivatives will also allow commercial use.

[Read the Commons Deed](#) | [View Legal Code](#)



[Choose by-sa license](#)

## Attribution (by)

This license lets others distribute, remix, tweak, and build upon your work, even commercially, as long as they credit you for the original creation. This is the most accommodating of licenses offered, in terms of what others can do with your works licensed under Attribution.

[Read the Commons Deed](#) | [View Legal Code](#)



[Choose by license](#)

[Opportunities](#)

[Press Kit](#)

[Press Releases](#)

[Newsletter](#)

[Events](#)

**The Commons**

[CC Network](#)

[Science Commons](#)

[ccInternational](#)

[ccLearn](#)

# International

Creative Commons works to "port" the core Creative Commons Licenses to different copyright legislations around the world. The porting process involves both linguistically translating the licenses and legally adapting them to particular jurisdictions.

This work is led by CC General Counsel [Diane Peters](#) (email) with Project Manager [Michelle Thorne](#) (email) and volunteer teams in each jurisdiction who are committed to introducing CC to their country and who consult extensively with members of the public and key stakeholders in an effort to adapt the CC licenses to their jurisdiction.

A complete overview of the porting process can be found at [http://wiki.creativecommons.org/International\\_Overview](http://wiki.creativecommons.org/International_Overview).

## Completed Licenses

We have completed the process and developed licenses for the following jurisdictions:



Argentina



Australia



Austria



Belgium



Brazil



Bulgaria



Canada



Chile

China  
Mainland

Colombia



Croatia

Czech  
Republic

Denmark



Ecuador



Finland



France



Germany



Greece



Guatemala



Hong Kong



Hungary



India



Israel



Italy



Japan

## Upcoming Project Jurisdictions

**Bangladesh:** Russell John (President) and Hasina Akhtar (Legal Advisor), Bangladesh Linux Users Alliance.

**Costa Rica:** Rolando Coto, Carlos E. Saborio Romero, Andrés Guadamuz, and Denis Campos; University of Costa Rica

**Egypt:** Hala Essalmawi; Bibliotheca Alexandrina

**Estonia:** Ene Tammeoru, Estonian Information Technology Foundation (EITF)

**Iceland:** Tryggvi Björgvinsson and Harald Gunnar Halldórsson; Reykjavik University and Félag um Stafrænt Frelsi á Íslandi (FSFI)

**Indonesia:** Ivan Lanin and Ari Juliano; Wikimedia Indonesia (WMID)

**Slovak Republic:** Martin Mojzis and Zuzana Adamova; Faculty of Law at Comenius University in Bratislava

**Tanzania:** Paul Kihwelo, Dean of Faculty, Open University of Tanzania.

**Tunisia:** Khaled Koubaa, Maître Nafaa Laribi, and Rafik

Search site

Go

[License your work](#)
[Find licensed works](#)
[The Licenses](#)
[Newsletter](#)
[Events](#)
[Supporters](#)
[Jobs](#)
[Developers](#)
[Case Studies](#)
[CC Talks With...](#)
[Education](#)
[Science](#)
[Contact](#)



# XML, OpenAPI, mashup

In web development, a **mashup** is a **web page or application that uses or combines data or functionality from two or many more external sources to create a new service.**

It implies easy, fast integration, frequently using open APIs and data sources to produce enriching results that were not necessarily the original reason for producing the raw source data.

To be able to permanently access the data of other services, mashups are generally client applications or hosted online. In the past years, more and more web applications provide APIs that enable software developers to easily integrate data and functions instead of building it themselves. Mashups can be considered to have an active role in the evolution of social software and Web 2.0.

**The term *mashup* is also used to describe a *remix of digital data*.**

<[http://en.wikipedia.org/wiki/Mashup\\_%28web\\_application\\_hybrid%29](http://en.wikipedia.org/wiki/Mashup_%28web_application_hybrid%29)>

**Esempi:**

<<http://www.earthalbum.com/>>; <<http://loc.alize.us/>>.

# Web 2.0: un video

*The Machine is Us/ing Us*

<<http://www.youtube.com/watch?v=6gmP4nk0EOE>>

# Alcuni esempi

**facebook**

<http://www.librarything.com/>

**Anobii**

<http://www.anobii.com/>

**Youtube**

<http://www.anobii.com/>

**Zopa**

<http://www.zopa.uk>

**ebay**

<http://www.ebay.com>

**Kiva**

<http://www.kiva.org/>

**Myspace**

<http://www.myspace.com>

**Twitter**

<http://twitter.com/>

**LinkedIn**

<http://www.linkedin.com/>

**Digg**

<http://digg.com/>

**LibraryThing**

<http://www.librarything.com/>

**Lastfm**

<http://www.last.fm/?setlang=en>



# Social network per accademici

**delicious**

<http://delicious.com/>

**Academia.edu**

<http://www.academia.edu>

**Connotea**

<http://www.connotea.org/>

**ResearchGate**

<http://www.researchgate.net/>

**CiteULike**

<http://citeulike.org/>

**Mendeley**

<http://www.mendeley.com/>

**Zotero**

<http://www.zotero.org>

**Altri?**

**MediaCommons**

<http://mediacommons.futureofthebook.org>

Citeulike

**citeulike** sponsored by  Springer[Browse](#) | [FAQ](#) | [News](#) | [Log in](#)

Search citeulike

**citeulike** is a free service for managing  
and discovering scholarly references

3,517,851 articles - 2,295 added today.

- Easily store references you find online
- Discover new articles and resources
- Automated article recommendations <sup>NEW</sup>
- Share references with your peers
- Find out who's reading what you're reading
- Store and search your PDFs

[Join now](#)



# Un account su Citeulike:

<http://www.citeulike.org/user/gioia>

# esercizio

1. Registratevi su Citeulike
2. aggiungete il bookmarklet alla barra degli strumenti
3. Postate un articolo/libro su CiteULike
  - a. da Amazon/JSTOR
  - c. manualmente

# Avete bisogno di aiuto?

Registrarsi a citeulike:

<http://www.youtube.com/watch?v=Yw8g6TYuvZ4>

Aggiungere il bookmarklet di CiteULike alla barra degli strumenti:

<http://www.youtube.com/watch?v=EEe1bCt-ux0>

Postare un articolo su CiteULike:

<http://www.youtube.com/watch?v=nurWqgl8l7o>



# Delicious

(social bookmarks)



Delicious

Search



Profile ▾

Add Link

Me

Feed

Add a Tag Filter

Normal View ▾



Make Public

Make Private

Delete

Add Tag

Rename Tag

Remove Tag

2 links

Recent



Delicious.com

11713 saves

<http://delicious.com/help/bookmarklets>

browser

delicious

bookmarklet



my page

1 save

<http://www.sp.unipi.it/hp/didonato>**ederina**

Joined 4 Minutes Ago

<http://edernet.wordpress.com/>RSS: [Public](#) | [Private](#)

Following: 0 Users

Followed by 0 Users

• 2 Links

Tweet 0

TAG BUNDLES

[Edit](#)

TAGS

By Count ▾

# Zotero:

collect

manage

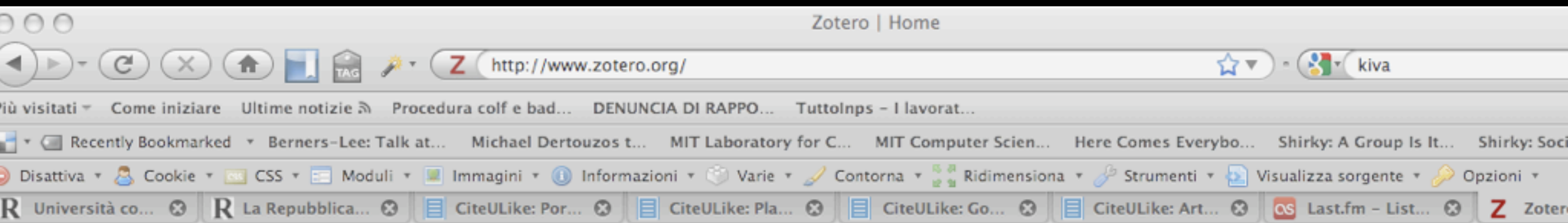
cite

and share

your research sources



# <http://www.zotero.org/>



# zotero

See it. Save it. Sort it. Search it. Cite it.

[Register](#) | [Find People](#) | [Login](#)

[add more file storage](#)

[Donate](#)

[Home](#)

[Groups](#)

[People](#)

[Support](#)

[Get Involved](#)

[Search](#)

**Zotero** [zoh-TAIR-oh] is a free, easy-to-use **Firefox extension** to help you **collect, manage, cite, and share** your research sources. It lives right where you do your work—in the **web browser** itself.



**Collect**



**Organize**



**Cite**



**Sync**



**Collaborate**

## Collect Everything

Record your online research with one click.

[download](#)

Latest version: **2.0.2**

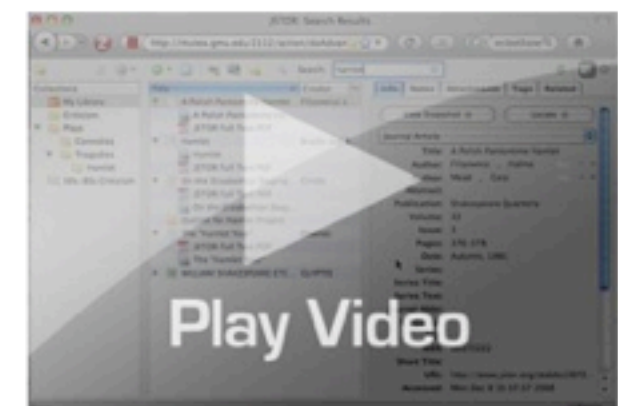
[Download word processor plugins](#)

[Download previous Zotero version \(1.0.10\)](#)

[What's new in Zotero 2.0?](#)

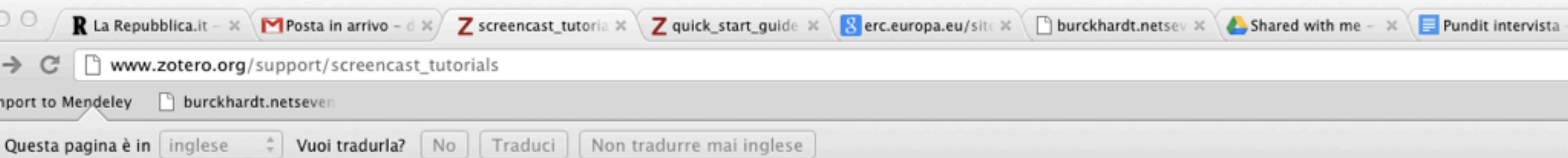
[Upgrading from Zotero 1.0](#)

[Having trouble installing?](#)



[Zotero News](#)

# [http://www.zotero.org/support/screencast\\_tutorials](http://www.zotero.org/support/screencast_tutorials)



# zotero

[Log In](#) · [Regi](#)[Upgrade Storage](#)[Home](#)[Groups](#)[People](#)[Documentation](#)[Forums](#)[Get Involved](#)[start](#) > [screencast tutorials](#)

Translations of this page: [en](#) [ar](#) [cs](#) [da](#) [de](#) [es](#) [fr](#) [hu](#) [id](#) [it](#) [ja](#) [ko](#) [nl](#) [pl](#) [pt](#) [ru](#) [sl](#) [sv](#) [tr](#) [zh](#)

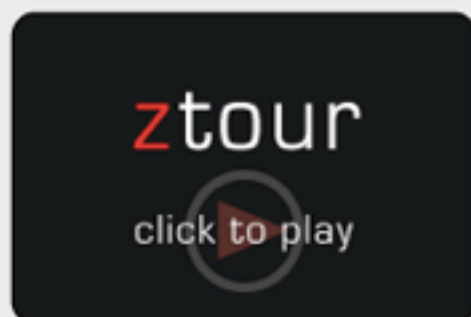
screencast\_tutorials.txt · Last modified:  
2012/04/12 16:03 by rmzelle

In an effort to make Zotero as user friendly as possible we have developed these screencasts demonstrating many of the basic functions of Zotero. Click on any of the images to watch screencasts detailing the features named below. Special thanks to Steve Bailey from CU-Boulder for preparing the introductory demo.

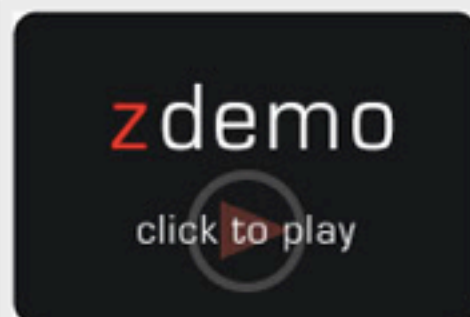
[Old revisions](#)

## Getting Started

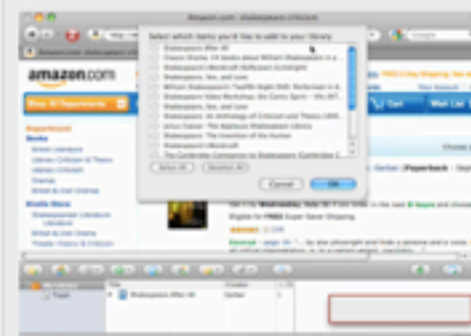
### Tour of Zotero



### Quick Bibliographic Demo



### Getting Stuff Into Zotero



## Finding Items

# Mendeley



MENDELEY

Sign up & Download

Sign In

Get Mendeley

What is Mendeley?

Papers

Groups

Papers

Search...



## It's time to change the way we do research.

Mendeley is a free **reference manager** and **academic social network** that can help you organize your research, collaborate with others online, and discover the latest research.

- Automatically **generate bibliographies**
- **Collaborate easily** with other researchers online
- Easily **import papers** from other research software
- **Find relevant papers** based on what you're reading
- Access your papers from **anywhere online**
- Read papers on the go, with our new **iPhone app**
- [View more features...](#)

First name

Last name

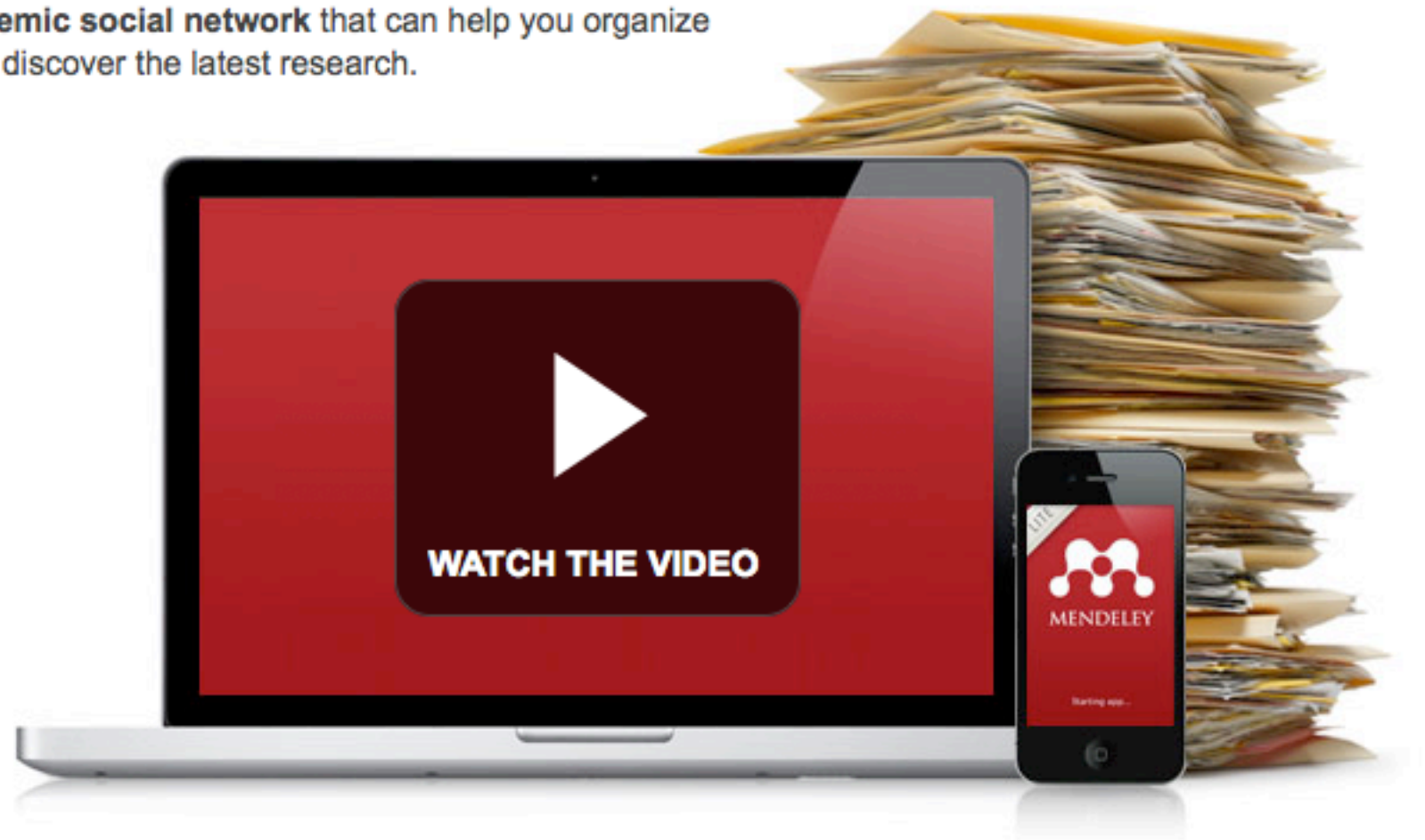
E-mail address

Sign up & Download

...or sign in with



Facebook



Overview

Discover Research

iPhone / iPad



# Mendeley

Vediamo come funziona:

<http://www.mendeley.com/features/>

<http://www.mendeley.com/how-we-help/>

<http://www.mendeley.com/videos-tutorials/>

# ResearchGate

ResearchGate

Already a member? [Log in](#) ▼

**ResearchGate was built for scientists, by scientists, with the idea that science can do more when it's driven by collaboration.**

**ResearchGate** began when two researchers discovered first-hand that working with a friend or colleague based on the other side of the world was no easy task. The rapid evolution of technology has opened the door to change; by providing you with the right tools, we strive to facilitate scientific collaboration on a global scale.

## **A hub for your research, your expertise, your contacts**

Research doesn't have to be confined to the lab; present yourself and your work, grow your contacts and connect with a global scientific community. All from your ResearchGate profile.

## **What, when, and where**

Staying on top of everything can be a challenge; we list thousands of scientific conferences from across the world, making it easy for you to stay informed, engaged and involved.

# ResearchGate

## A place to get answers

Progress always involves obstacles; we enable you to discuss your work and methodology, help others and find solutions together.

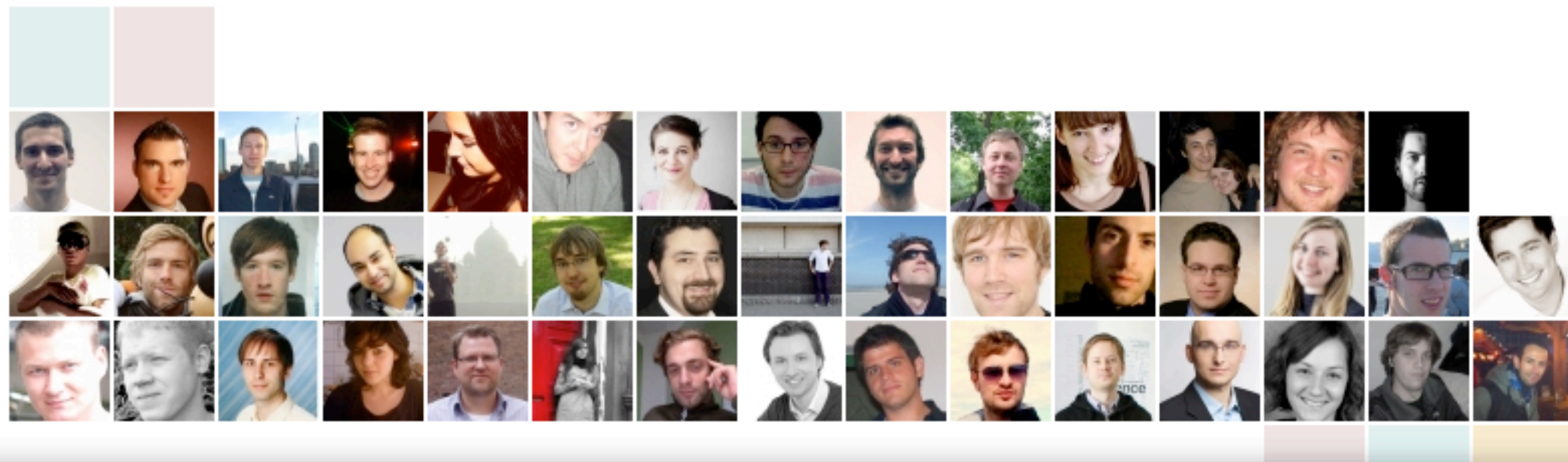
## Millions of publications at your fingertips

It can be tricky to get hold of the right publications; so whether you want to find something specific, discover something new or share your own work with the community, we have the tools to make it happen.

## The right person, the right career

Jobs aren't always easy to come by; our specialised listings devoted to science and research make spotting your ideal position a swift and straightforward process. It's that simple.

## Meet the team

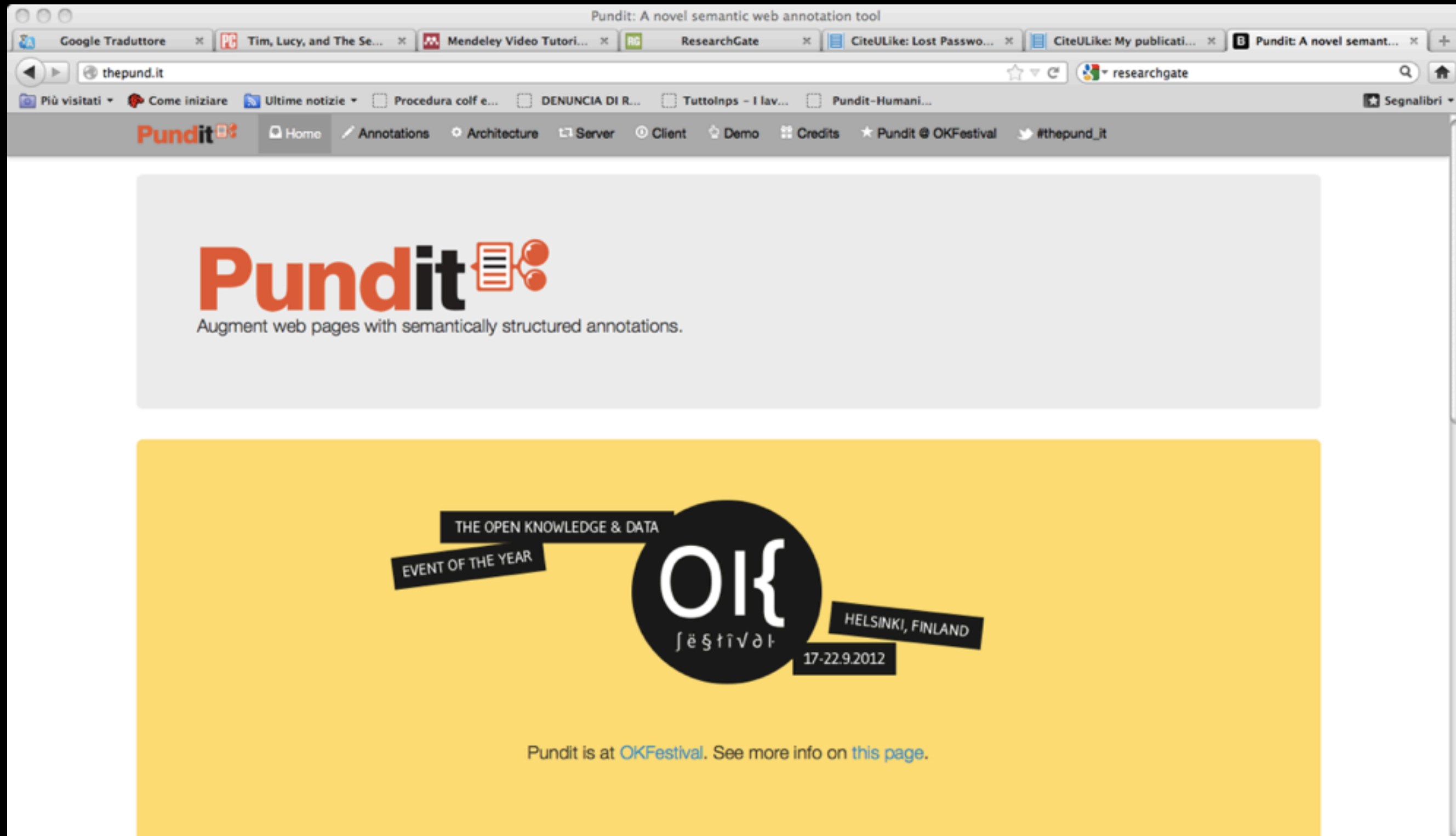


*ResearchGate is the professional network for scientists and researchers. Join today, it's free!*

**Sign Up »**



# Pundit: aggiungere “semantica” alla conoscenza



### 3. Disseminare i propri risultati

# In ambito accademico disseminare significa pubblicare

**Pubblicare** vuol dire rendere le produzioni intellettuali accessibili al pubblico dei lettori

**Due** sistemi di pubblicazione accademica:

1. L'età della stampa

2. L'età del Web



# 1: l'età della stampa

## Attori

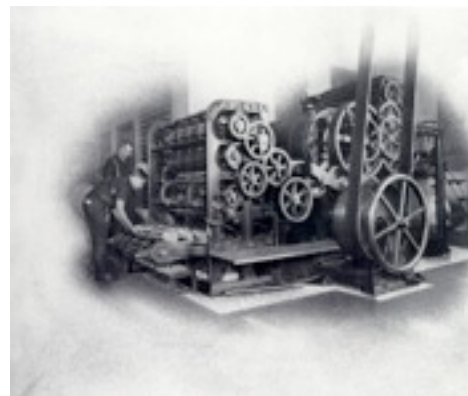
Istituzioni scientifiche  
(Università)



Scienziati



Editori



Bibliotecari



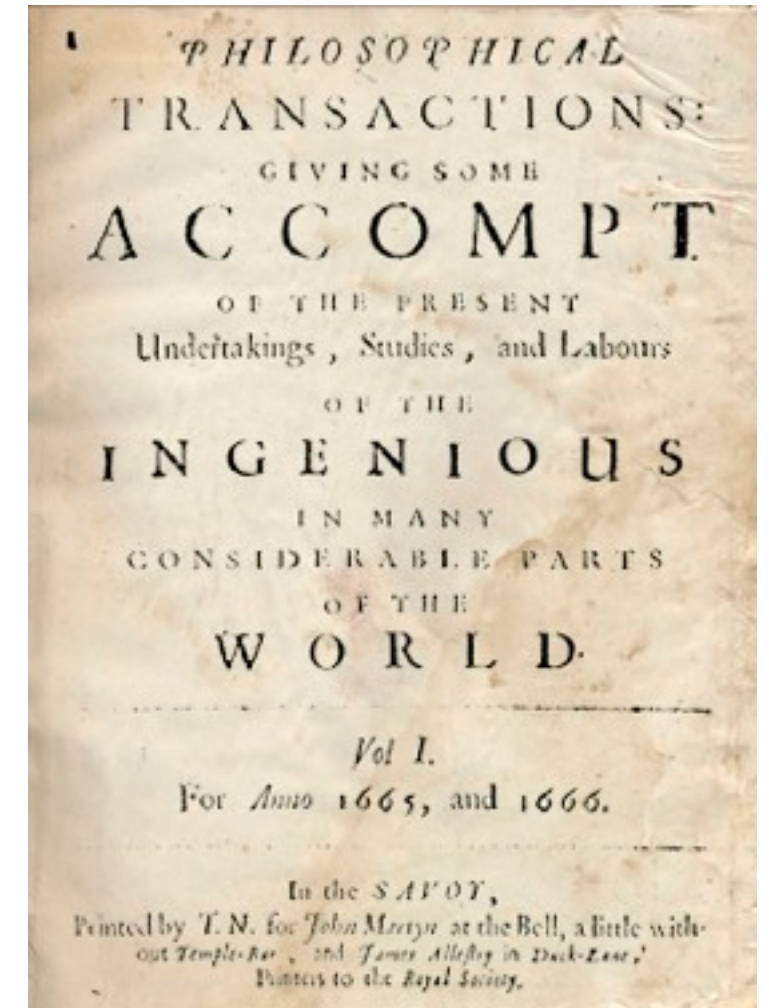
# Contesto tradizionale

Londra, XVII secolo

- Nascita delle riviste scientifiche (disseminazione)
- Nascita del peer reviewing (procedure di valutazione)
- Nascita del copyright (contesto giuridico)

# l'età della stampa > all'inizio

- **1665: *Philosophical Transactions***
- **Peer reviewing**
- @Royal Society
- Perusal (per ottenere una fellowship)
- Un lavoro poteva poi essere o meno stampato
- **1710 Statute of Anne**
- Prima legge sul copyright (termina il sistema dei privilegi)
- 14 anni (+14)



## Cfr:

- J.C. Guéron, *In Oldenburg's Long Shadow: Librarians, Research Scientists, Publishers, and the Control of Scientific Publishing*, Association of Research Libraries, Proceedings of the 138th Annual Meeting (2001).  
<<http://www.arl.org/arl/proceedings/138/guedon.html>>
- F. Di Donato, *La scienza e la rete*, FUP, 2009, on-line at: <<http://www.fupress.com/scheda.asp?IDV=1953>>



# età della stampa > dopo la seconda guerra mondiale

- (anni 30: Legge di Bradford e “core journals”)
- 1960: Garfield e lo ***Science Citation Index***: strumento finalizzato a permettere alle biblioteche di scegliere la letteratura di rilievo, ai fini delle loro politiche di acquisto; e pensato perché i ricercatori potessero costruire facilmente bibliografie pertinenti e complete.

Garfield pone al centro della riflessione due elementi essenziali per la comunicazione scientifica moderna: l'importanza delle **citazioni**; e il problema della **selezione dell'informazione scientifica**.



# SCI e Impact Factor

Lo SCI si sarebbe rivelato molto utile nelle scienze umane, filosofiche e sociali, per tracciare percorsi che ricostruissero l'origine e l'evoluzione di un'idea e nelle ricerche storiche al fine di valutare il significato di un'opera calcolandone **impatto** sulla letteratura e sul pensiero successivi.

Un tale **fattore d'impatto (Impact Factor)** avrebbe potuto rivelarsi uno strumento di valutazione più indicativo e significativo rispetto al conteggio assoluto del numero delle pubblicazioni scientifiche.



# Bibliometria

Scienza applicata a vari ambiti scientifici che utilizza tecniche matematiche e [statistiche](#) per analizzare i modelli di distribuzione delle pubblicazioni scientifiche e per verificarne il loro impatto all'interno delle comunità scientifiche. Questa scienza, che comprende l'analisi bibliometrica, è alquanto recente: essa infatti nasce negli [anni venti](#) e si è sviluppata enormemente in conseguenza della disponibilità online di banche dati di grandi dimensioni. La bibliometria si inserisce all'interno dell'ambito di studi più ampio, definito [scientometria](#), ossia la scienza per la misura e l'analisi della scienza.

I primi lavori di bibliometria e scientometria risalgono ai primi decenni del [XX secolo](#). [Alfred Lotka](#) , [Samuel Bradford](#) e [George Kingsley Zipf](#) ne elaborarono i fondamenti e le leggi tra il [1926](#) e il [1935](#). Negli [anni cinquanta](#) [John Derek de Solla Price](#) elaborò le basi di una disciplina basata sul ricorso agli articoli scientifici come indicatori quantitativi dell'attività di ricerca[\[1\]](#). Nello stesso periodo [Eugene Garfield](#) sviluppò l'idea di usare le citazioni bibliografiche contenute negli articoli scientifici per valutarne l'[impatto](#)[\[2\]](#)[\[3\]](#); il primo volume di Science Citation Index apparve nel [1963](#).

<<http://it.wikipedia.org/wiki/Bibliometria>>



# Ma come funziona l'IF?

Lo SCI, «da sistema per risolvere al meglio un problema organizzativo all'interno delle biblioteche (la possibilità di individuare con certezza quali periodici acquistare, in base al loro reale utilizzo) divenne un sistema di valutazione della scienza nel suo complesso», soprattutto grazie all'adozione dell'Impact Factor (IF), una forma di misura standardizzata introdotta dall'Institute of Scientific Information (ISI) che si calcola dividendo il numero di citazioni che una rivista riceve, in un dato anno, relative agli articoli pubblicati nei due anni precedenti, per il numero di articoli pubblicati in quegli stessi anni.

Ad esempio, il fattore d'impatto di una rivista per il 2008 è il numero di citazioni ricevute nel 2008 relative agli articoli pubblicati nel 2007 e nel 2006, diviso per il numero di articoli pubblicati nel 2007 e nel 2006.



# Limiti dell'IF

Contare le citazioni di una rivista non può essere considerato un criterio per valutare la qualità di un articolo. Inoltre, come osserva Alessandro Figà-Talamanca

“altro è che un lavoro sia citato con una frase del tipo: “Il presente lavoro si basa sul lavoro fondamentale e pionieristico di Tizio”. Altro che invece sia citato con una frase che dice: “I risultati su questo problema ottenuti da Caio non sono affidabili perché basati su un’errata metodologia”. Altro ancora è che Sempronio appaia citato tra decine di altri autori come uno dei tanti ricercatori che hanno studiato lo stesso problema. Nel primo e nel secondo caso naturalmente non si può nemmeno prendere per oro colato il giudizio di chi cita il lavoro. Nel terzo caso (che è quello ormai prevalente) l’informazione fornita dalla citazione, in merito al lavoro citato, è pressoché nulla. In molti casi non si può nemmeno dire che chi cita abbia letto l’articolo citato, meno che mai che ne sia stato influenzato. Parlare di “impatto” sulla ricerca scientifica dei risultati del lavoro citato è, in questo caso, del tutto arbitrario”.

Inoltre il sistema SCI-IF, di proprietà dello Institute for Scientific Information (un'azienda privata che dal 1992 appartiene alla Thomson Reuters Corporation), pone un problema di democraticità nell'accesso all'informazione.



# Effetti del sistema SCI-IF

Il catalogo, nato in origine per scopi bibliografici, ha finito per contribuire a produrre politiche dei prezzi di tipo monopolistico tipiche di un mercato anelastico, con pesanti conseguenze: proprio a causa del loro ruolo decisivo nella valutazione del fattore d'impatto, le biblioteche universitarie non possono non abbonarsi alle riviste ISI-SCI, quale che sia il costo di tali abbonamenti.

La scarsa elasticità della domanda e il carattere oligopolistico delle multinazionali dell'editoria scientifica hanno reso possibile un aumento sproporzionato dei prezzi di abbonamento delle riviste.



# altri indici citazionali

**H-index (Index of Hirsch):** is an index that attempts to measure both the productivity and impact of the published work of a scientist or scholar.

A scientist has index  $h$  if  $h$  of his/her  $N_p$  papers have at least  $h$  citations each, and the other  $(N_p - h)$  papers have no more than  $h$  citations each.

In other words, a scholar with an index of  $h$  has published  $h$  papers each of which has been cited in other papers at least  $h$  times. Thus, the  $h$ -index reflects both the number of publications and the number of citations per publication.

# webometria e valutazione della scienza

## Thomson-Reuters, Web of Science:

### DB: Web of science:

>10,000 Journals (including Core Journals)  
> 110,000 conference proceedings,  
Universities pay very high subscriptions rates  
It offers additional services such as ResearcherID, EndNote.

### Index: IF:

In a given year, the impact factor of a journal is the average number of citations to those papers that were published during the two preceding years.

For example, the 2008 impact factor of a journal would be calculated as follows:

$A$  = the number of times articles published in 2006 and 2007 were cited by indexed journals during 2008  
 $B$  = the total number of "citable items" published in 2006 and 2007. ("Citable items" are usually articles, reviews, proceedings, or notes; not editorials or Letters-to-the-Editor.)

2008 impact factor =  $A/B$

**openness:** in entrance/ in exit: no/no

# SCOPUS, Elsevier:

## DB: SCOPUS

Nearly **18,000 titles** from more than **5,000 international publishers**, including coverage of

- 16,500 peer-reviewed journals (inc > 1,200 Open Access journals)
- 600 trade publications
- 350 book series
- Extensive conference coverage (3.6 million conference papers)

40 million records, of which:

- 20 million records going back to 1996 seeking to capture complete metadata (78% include references)
- 20 million pre-1996 records captured without references go back as far as 1823.

## Index: h index

The ***h-index*** is an [index](#) that attempts to measure both the scientific productivity and the apparent scientific impact of a scientist.

The index is based on the distribution of [citations](#) received by a given researcher's publications. Hirsch writes:

*A scientist has index  $h$  if  $h$  of [his/her]  $N_p$  papers have at least  $h$  citations each, and the other  $(N_p - h)$  papers have at most  $h$  citations each.*

**openness:** in entrance/exit: ?no/no



# la webometria e i grandi editori

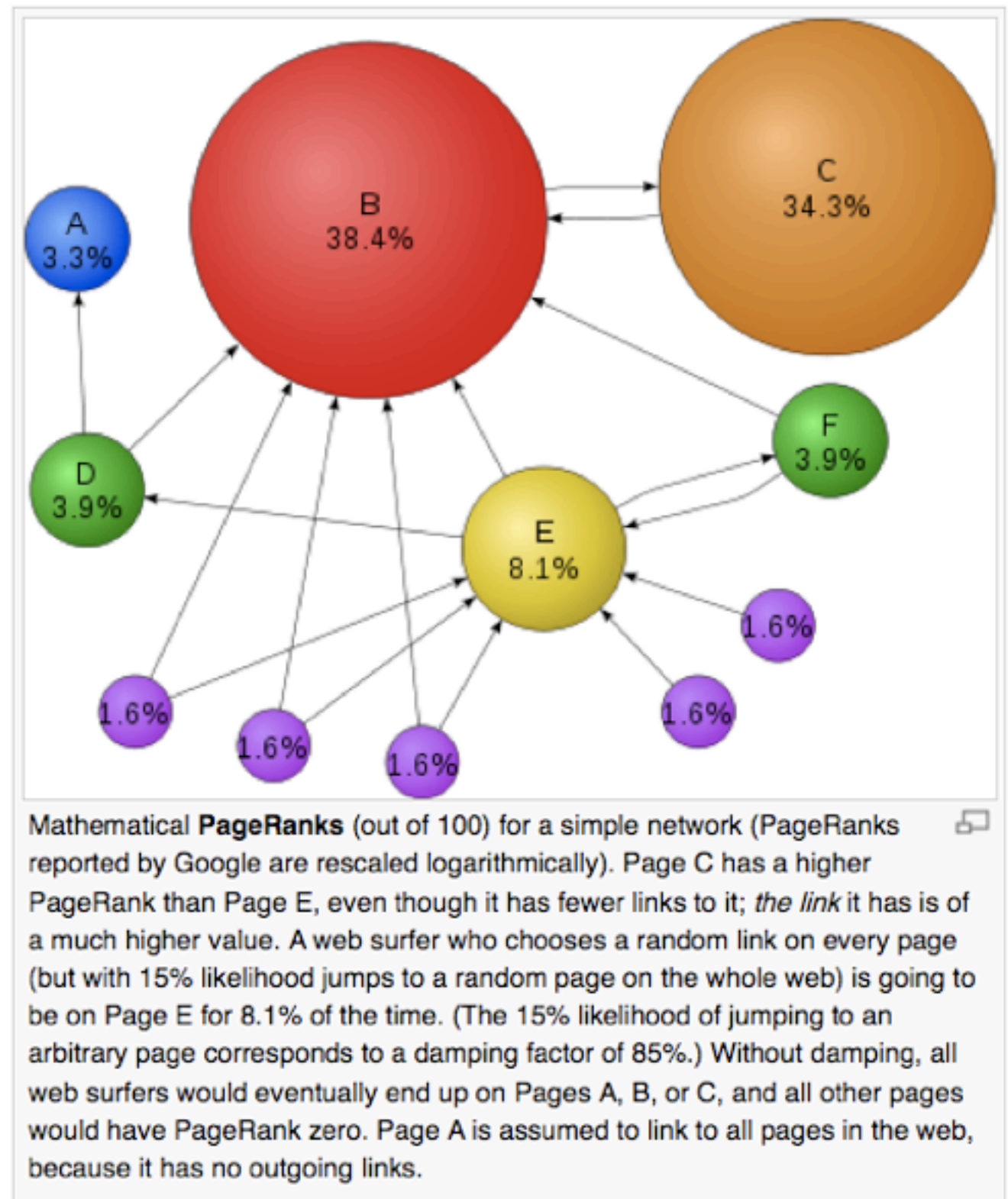
**Google scholar:**

**DB: google indexed pages**

**Index: Pagerank**

**PageRank** is a [link analysis](#) algorithm used by the [Google](#) Internet [search engine](#) that assigns a numerical weighting to each element of a [hyperlinked set](#) of documents, such as the [World Wide Web](#), with the purpose of "measuring" its relative importance within the set.

**openness:** entrance/exit: no/yes



# un sistema “publish or perish”

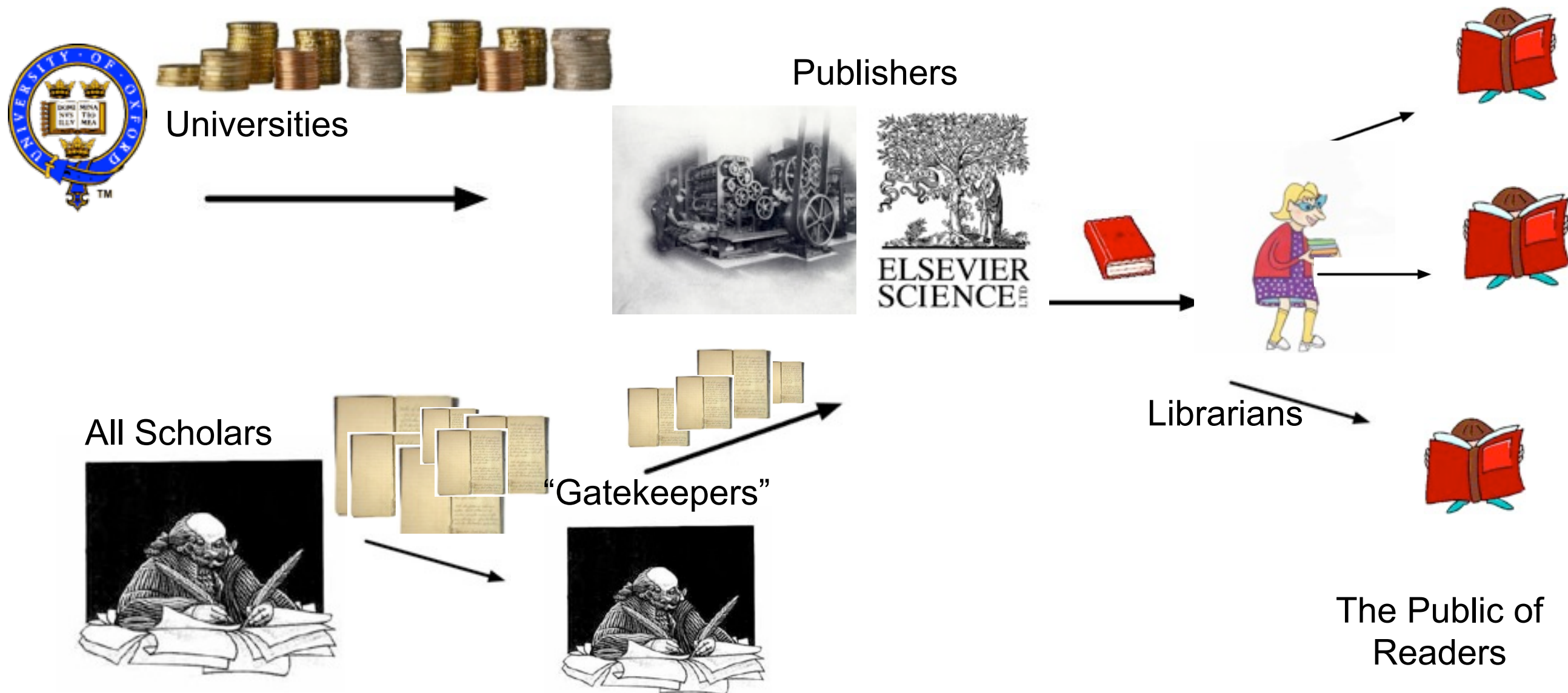
gli accademici sono obbligati a pubblicare  
(per ottenere finanziamenti, per  
reclutamento e carriera)

pratica del “salami publishing”

la quantità prevale sulla qualità

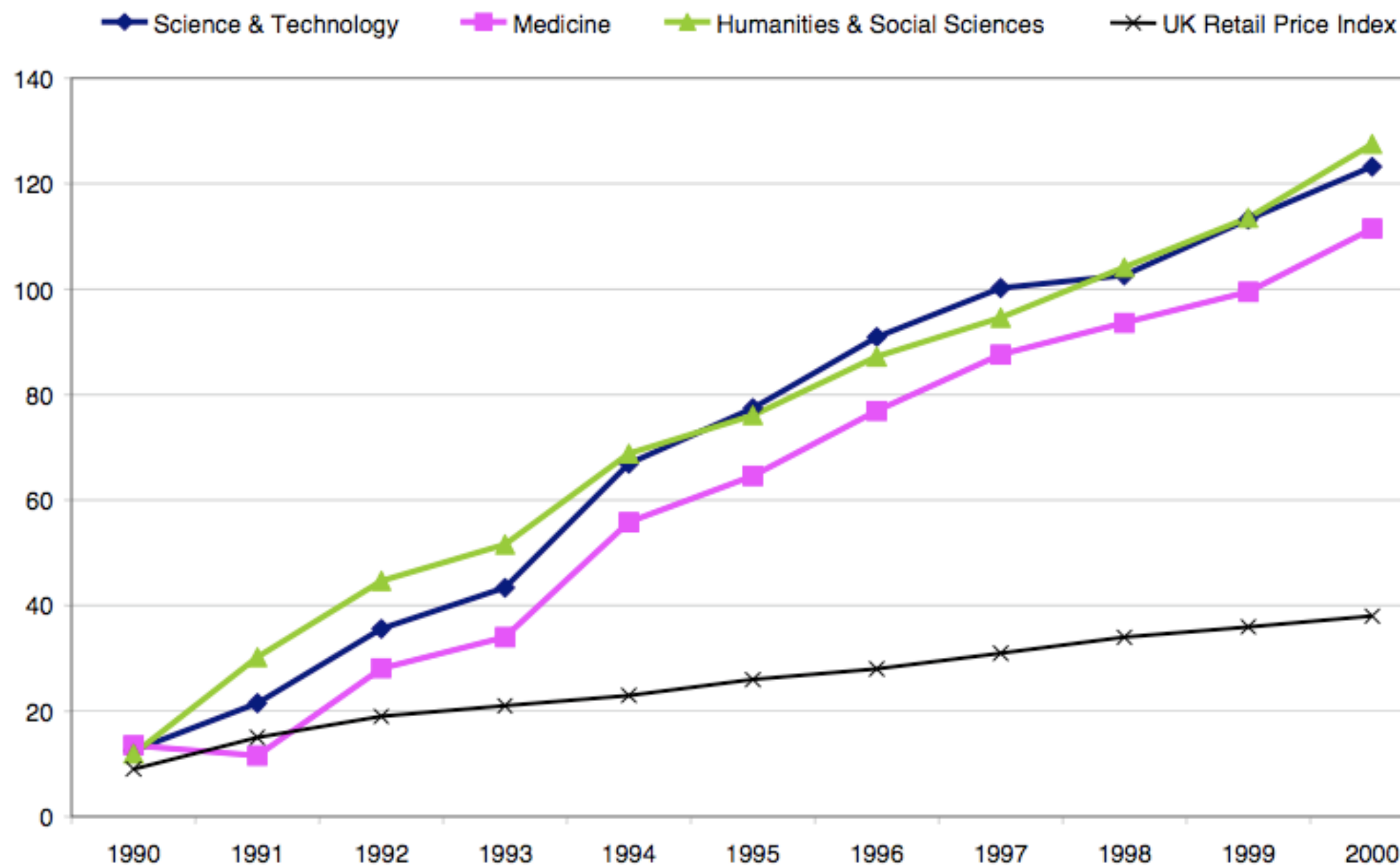
# età della stampa > il mercato

- Mercato anelastico -----> “crisi del prezzo dei periodici”





**Figure 7. Journal price increases, 1990-2000**  
(cumulative annual percentage change)



Source: Blackwell's Periodical Prices. Cited by Department of Trade and Industry (2002), *Publishing in the Knowledge Economy: Competitiveness analysis of the UK publishing media sector*, Department of Trade and Industry, London, p37.

esplorando  
alternative

# 2.

## l'età del web

### Attori

editori



Istituzioni scientifiche



Scienziati



Bibliotecari

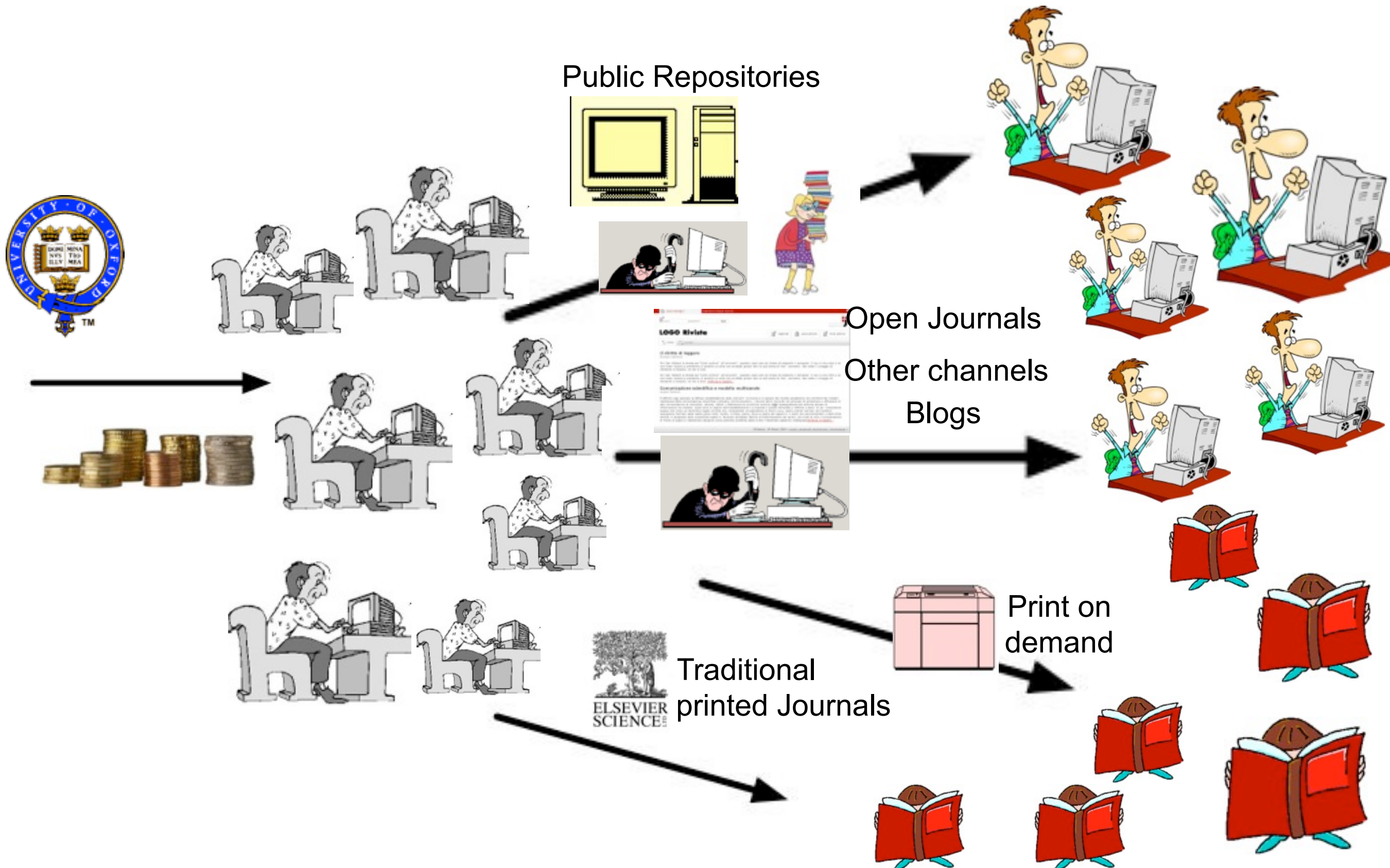


Informatici





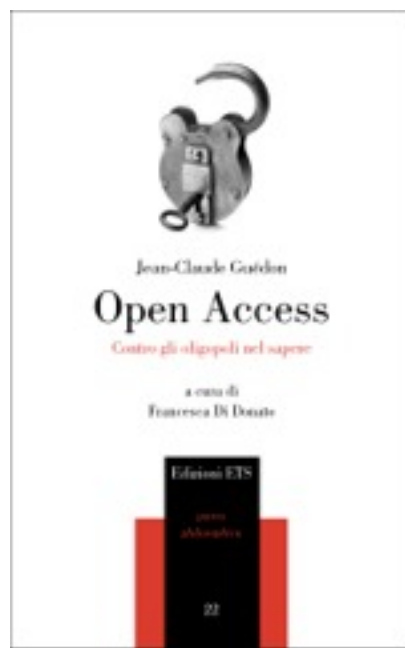
# Un nuovo scenario



# Open Access

un modello di comunicazione  
e di valutazione  
della scienza

# Perché?



Dal secondo dopoguerra il sistema di comunicazione (e di valutazione) nella scienza si è trasformato.

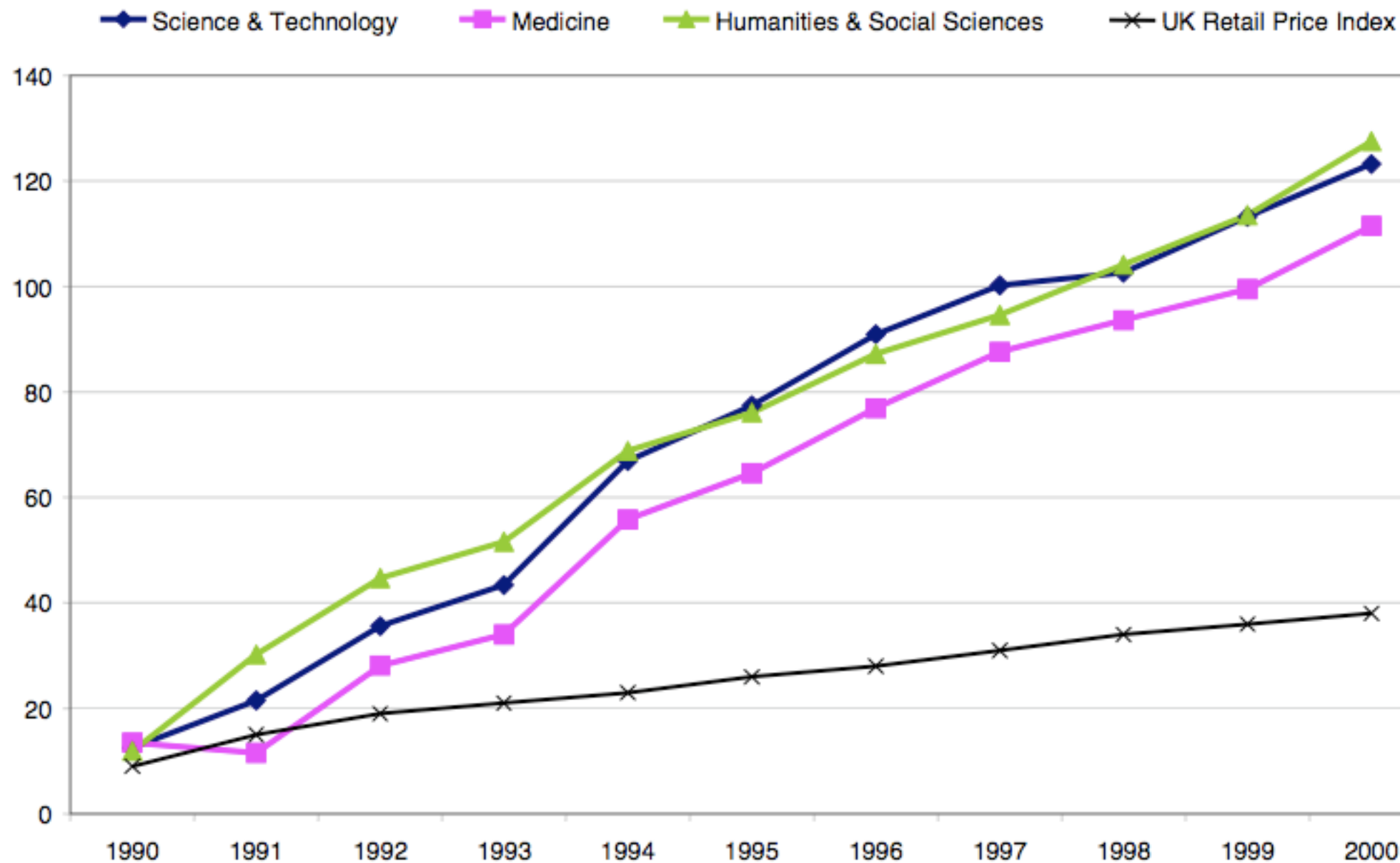
- aumento delle riviste
- indice delle citazioni: SCI-Web of Science di ISI + Impact Factor
- grande potere ad alcuni editori commerciali, che possiedono le riviste ISI
- sistema anglo-centrico
- divisione artificiosa tra scienza mainstream e scienza periferica.

“Crisi del prezzo dei periodici”:

- tra il 1975 e il 1995, i costi degli abbonamenti alle riviste fondamentali (quelle ISI) sono cresciuti del 300% oltre l'inflazione
- questo fenomeno ha messo in ginocchio le biblioteche universitarie e condiziona pesantemente l'impiego dei fondi di ricerca da parte dei ricercatori



**Figure 7. Journal price increases, 1990-2000**  
(cumulative annual percentage change)



Source: Blackwell's Periodical Prices. Cited by Department of Trade and Industry (2002), *Publishing in the Knowledge Economy: Competitiveness analysis of the UK publishing media sector*, Department of Trade and Industry, London, p37.

L'open access  
nasce negli anni 90  
come risposta  
alla crisi  
del prezzo dei periodici

# Una definizione



“La **letteratura** ad accesso aperto (open access) è:

**digitale, online, gratuita e libera da alcune restrizioni** dettate dalle licenze per i **diritti di sfruttamento commerciale**.

Queste condizioni sono possibili grazie ad **Internet** e al consenso dell'autore o del titolare dei **diritti d'autore**.”



# due “canali”....

## Self-Archiving (*green road*)

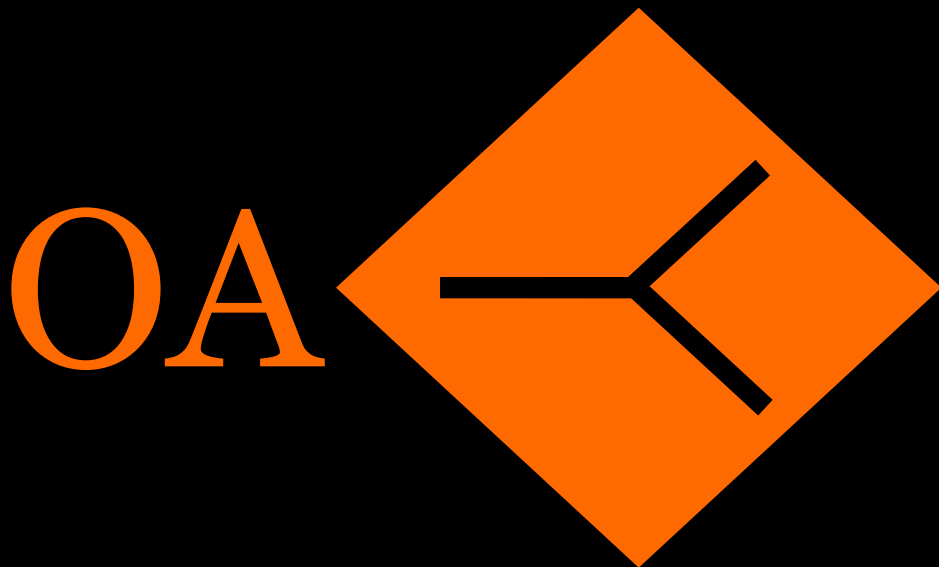
“Gli **archivi aperti o repositories** non effettuano referaggio ma **rendono** semplicemente i loro contenuti **liberamente disponibili** al mondo, siano essi **preprints non referati** o **postprints referati**, o entrambi.

Gli archivi **possono essere istituzionali**, come quelli di università e centri di ricerca, **o disciplinari**, come quelli per la fisica o l'economia. ...”

## OA publishing (*gold road*)

“Le **riviste ad accesso aperto** effettuano il **referaggio** e rendono i loro articoli selezionati liberamente disponibili al mondo. I loro costi consistono nella gestione dei processi editoriali, dal referaggio alla manipolazione dei manoscritti, allo spazio sul server. Le riviste ad accesso aperto coprono i propri costi in un modo molto simile alle radio e alle televisioni: coloro che hanno interesse alla disseminazione dei contenuti pagano anticipatamente i costi di produzione in modo che l'accesso alla lettura sia libero. ...”

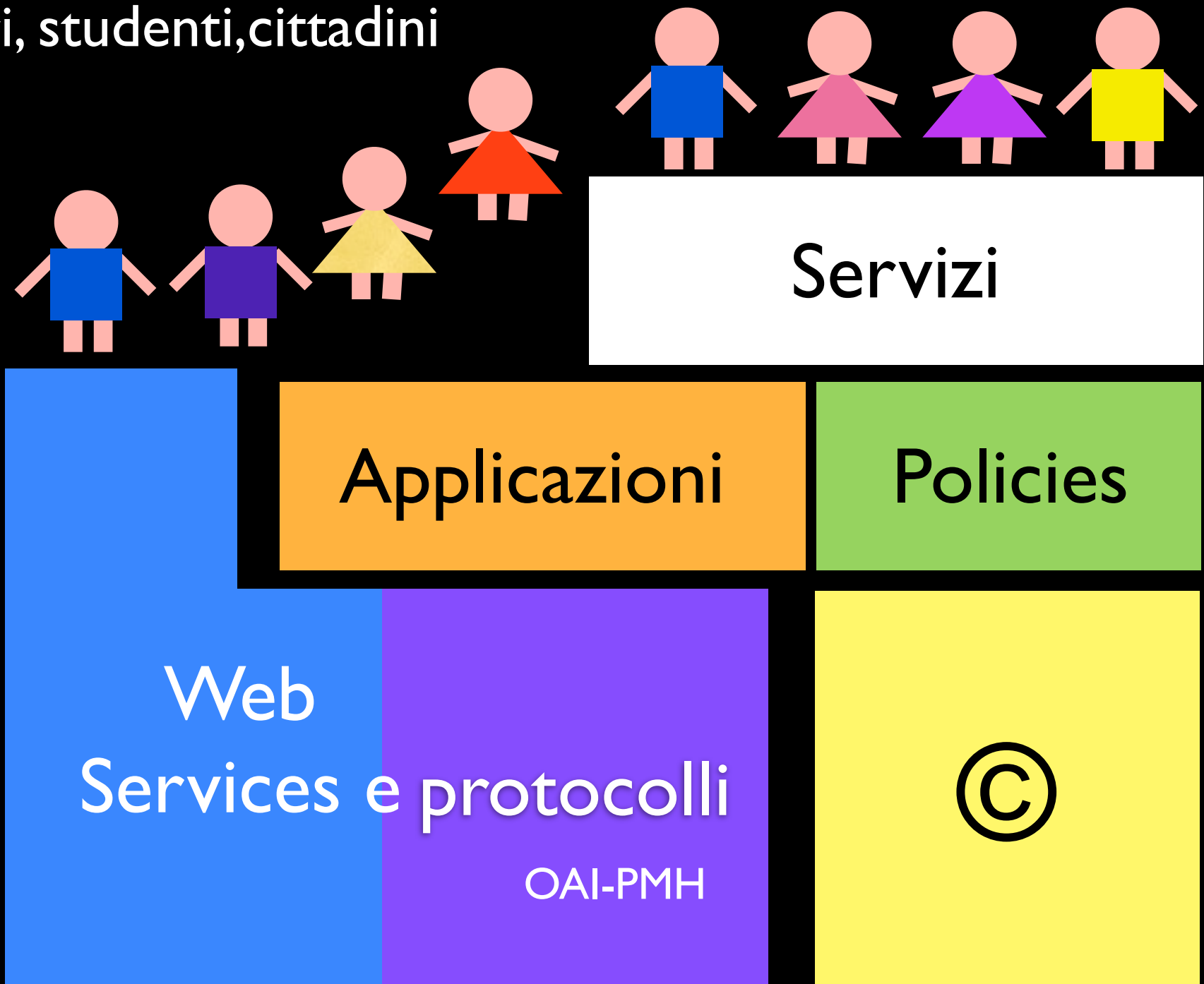
[Suber]



# OA “layers cake”

Istituzioni, ricercatori, studenti, cittadini

Documenti  
e  
dati



2010

2003

2000

1999

1991

Internet, Web, Web semantico

# Strumenti per l'OA

## Applicazioni per creare archivi e riviste

- Eprints
- DSPace
- CDSWare
- Fedora
- OJS
- HyperJournal
- .....

?

## Policies

Dichiarazioni internazionali  
Politiche istituzionali!!

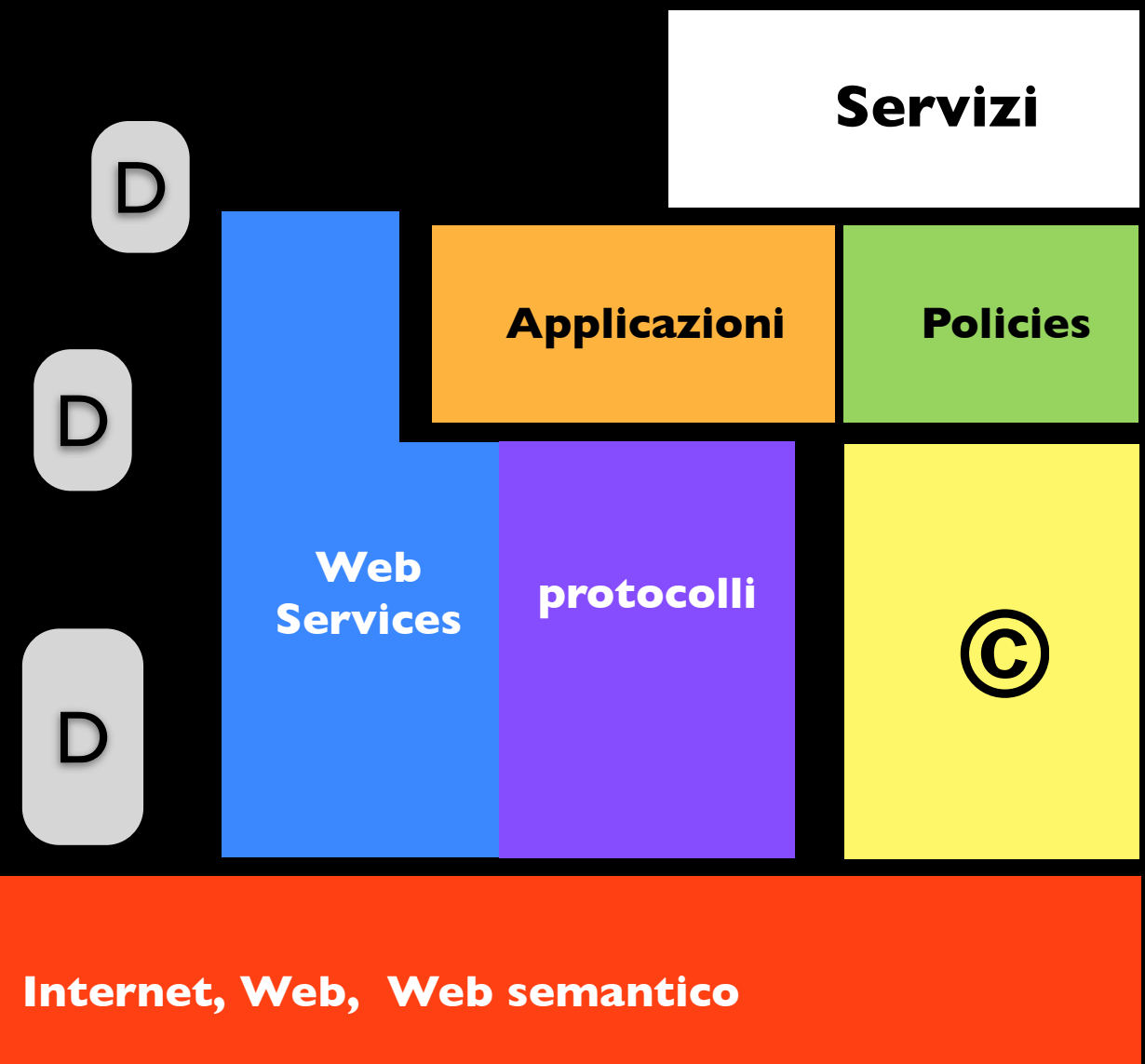
## Web services

Protocolli, standard  
OAI-PMH, RSS, Dublin Core, ...

## Diritti

Legge 22 aprile 1941, n. 633  
Licenze (CC)

## Infrastruttura





# Dichiarazione di Berlino (2003)

## **Accesso aperto alla letteratura scientifica (Dichiarazione di Berlino)**

### **Premessa**

Internet ha radicalmente modificato le realtà pratiche ed economiche della distribuzione del sapere scientifico e del patrimonio culturale. Per la prima volta nella storia, Internet offre oggi l'occasione di costituire un'istanza globale ed interattiva della conoscenza umana e dell'eredità culturale e di offrire la garanzia di un accesso universale.

Noi, i firmatari, ci impegniamo ad affrontare le sfide di Internet come mezzo funzionale emergente per la diffusione della conoscenza. Siamo certi che questi sviluppi saranno in grado di incidere significativamente tanto sulla natura delle pubblicazioni scientifiche quanto sul sistema esistente di valutazione della qualità scientifica.

In accordo con lo spirito della Dichiarazione della Budapest Open Access Initiative, la Carta di ECHO e il Bethesda Statement sull'Open Access Publishing, abbiamo redatto la Dichiarazione di Berlino per promuovere Internet quale strumento funzionale alla conoscenza scientifica generale di base e alla speculazione umana e per indicare le misure che le figure dominanti nelle politiche di ricerca, le istituzioni scientifiche, i finanziatori, le biblioteche, gli archivi ed i musei devono tenere in considerazione.

## **Obiettivi**

La nostra missione di disseminazione della conoscenza è incompleta se l'informazione non è resa largamente e prontamente disponibile alla società. Occorre sostenere nuove possibilità di disseminazione della conoscenza, non solo attraverso le modalità tradizionali ma anche e sempre più attraverso il paradigma dell'accesso aperto via Internet. Definiamo l'accesso aperto come una fonte estesa del sapere umano e del patrimonio culturale che siano stati validati dalla comunità scientifica.

Per mettere in pratica la visione di un'istanza globale ed accessibile del sapere, il Web del futuro dovrà essere sostenibile, interattivo e trasparente. I contenuti ed i mezzi di fruizione dovranno essere compatibili e ad accesso aperto.

## **Definizione di contributi ad accesso aperto**

Accreditare l'accesso aperto quale procedura meritevole richiede idealmente l'impegno attivo di ogni e ciascun produttore individuale di conoscenza scientifica e di ciascun depositario del patrimonio culturale. I contributi ad accesso aperto includono le pubblicazioni di risultati originali della ricerca scientifica, i dati grezzi e i metadati, le fonti, le rappresentazioni digitali grafiche e di immagini e i materiali multimediali scientifici.

Ciascun contributo ad accesso aperto deve soddisfare due requisiti:

1. L'autore(i) ed il detentore(i) dei diritti relativi a tale contributo garantiscono a tutti gli utilizzatori il diritto d'accesso gratuito, irrevocabile ed universale e l'autorizzazione a riprodurlo, utilizzarlo, distribuirlo, trasmetterlo e mostrarlo pubblicamente e a produrre e distribuire lavori da esso derivati in ogni formato digitale per ogni scopo responsabile, soggetto all'attribuzione autentica della paternità intellettuale (le pratiche della comunità scientifica manterranno i meccanismi in uso per imporre una corretta attribuzione ed un uso responsabile dei contributi resi pubblici come avviene attualmente), nonché il diritto di riprodurre una quantità limitata di copie stampate per il proprio uso personale.
2. Una versione completa del contributo e di tutti i materiali che lo corredano, inclusa una copia della autorizzazione come sopra indicato, in un formato elettronico secondo uno standard appropriato, è depositata (e dunque pubblicata) in almeno un archivio in linea che impieghi standard tecnici adeguati (come le definizioni degli Open Archives) e che sia supportato e mantenuto da un'istituzione accademica, una società scientifica, un'agenzia governativa o ogni altra organizzazione riconosciuta che persegua gli obiettivi dell'accesso aperto, della distribuzione illimitata, dell'interoperabilità e dell'archiviazione a lungo termine.



## **Sostenere la transizione verso il paradigma dell'accesso aperto elettronico**

Le nostre organizzazioni sono interessate all'ulteriore promozione del nuovo paradigma dell'accesso aperto per offrire il massimo beneficio alla scienza e alla società. Perciò intendiamo favorirne il progresso:

- incoraggiando i nostri ricercatori e beneficiari di finanziamenti per la ricerca a pubblicare i risultati del loro lavoro secondo i principi dell'accesso aperto
- incoraggiando i detentori del patrimonio culturale a supportare l'accesso aperto mettendo a disposizione le proprie risorse su Internet
- sviluppando i mezzi e i modi per valutare i contributi ad accesso aperto e le pubblicazioni in linea, così da preservare gli standard qualitativi della validazione e della buona pratica scientifica
- difendendo il riconoscimento delle pubblicazioni ad accesso aperto ai fini delle valutazioni per le promozioni e l'avanzamento delle carriere
- difendendo il merito intrinseco dei contributi ad un'infrastruttura ad accesso aperto attraverso lo sviluppo di strumenti di fruizione, la fornitura di contenuti, la creazione di metadati o la pubblicazione di articoli individuali.

Noi riconosciamo che il passaggio all'accesso aperto modifica la disseminazione della conoscenza nei suoi aspetti legali e finanziari. Le nostre organizzazioni mirano a trovare soluzioni che sostengano futuri sviluppi degli attuali inquadramenti legali e finanziarie al fine di facilitare l'accesso e l'uso ottimale.

[<oa.mpg.de/files/2010/04/BerlinDeclaration\\_it.pdf>](http://oa.mpg.de/files/2010/04/BerlinDeclaration_it.pdf)



# Dichiarazione di Messina (2004)

75 rettori di Università italiane su 77 sottoscrivono la Dichiarazione di Berlino, impegnandosi a sostenere l'accesso aperto.

L'Open Access ci spinge a  
ripensare il rapporto  
tra **il modo in cui produciamo i**  
**risultati di ricerca e come li**  
**pubblichiamo** (certifichiamo,  
archiviamo, valutiamo,  
disseminiamo)

# Un nuovo paradigma di comunicazione scientifica

Nel paradigma tradizionale, sono in pochi a decidere che cosa merita di essere letto: gli autori inviano il documento agli editor che lo trasmettono ai valutatori, e i cui giudizi vincolano gli editor nel decidere che cosa pubblicare. In un sistema più democratico, reso possibile dalle nuove tecnologie, **gli autori possono sottoporre i loro contributi a editor e valutatori, ma anche coinvolgere l'intera comunità scientifica in tale attività, fino a sottoporre i testi che producono direttamente al giudizio del pubblico, combinando criteri formali e informali.**



# Altre forme di peer review

Al peer review tradizionale si sono aggiunte nuove forme di valutazione qualitative.

1. Il **peer review ex post**, una forma di certificazione della qualità applicata a risultati già pubblicati in archivi (tipicamente sui preprint, cioè su articoli che non siano stati sottoposti a referaggio) tramite l'uso di overlay journal che selezionano il contenuto archiviato nei repository (generalmente in forma di preprint ovvero di articoli non referati) validandoli.

2. L'**open peer review**, un'altra forma di peer review a posteriori applicato a risultati già pubblicati e sottoposti a referaggio. Un esempio consiste nel lasciare la possibilità di commentare il testo ai lettori in modo simile a quanto avviene sui blog.

3. Il **soft peer review**, che consente di usare i metadati collaborativi archiviati su “online reference managers” come CiteUlike e Connotea, con la possibilità di gestire ampie quantità di dati, che sul Web sono già a disposizione. Nuovi strumenti e parametri di valutazione si basano sull'uso di metadati semantici, indicatori di popolarità, sistemi di rating assegnati dai lettori (con un procedimento analogo a quello che avviene sul sito Venere.com, che permette di vedere i giudizi che gli utenti hanno dato sugli alberghi disponibili) e annotazioni collaborative da parte degli utenti.

# OA e valutazione

L'OA viene sempre più spesso indicato come prerequisito/bonus in sede di valutazione nazionale ed europea.

Garantisce una valutazione trasparente e più completa:

- indice h (citazioni)
- indici di download
- posizionamento su google, google scholar
- indici basati su social network analysis (considerano la posizione dell'articolo nel network, invece si basarsi solo sulle citazioni ricevute)
- trattamento semantico dei dati (vantaggi dell'open data, non solo per la valutazione..)
- ?

## **Pluralità**

di fonti e di criteri

(non solo indici: peer-to-peer review!)

## **Trasparenza**

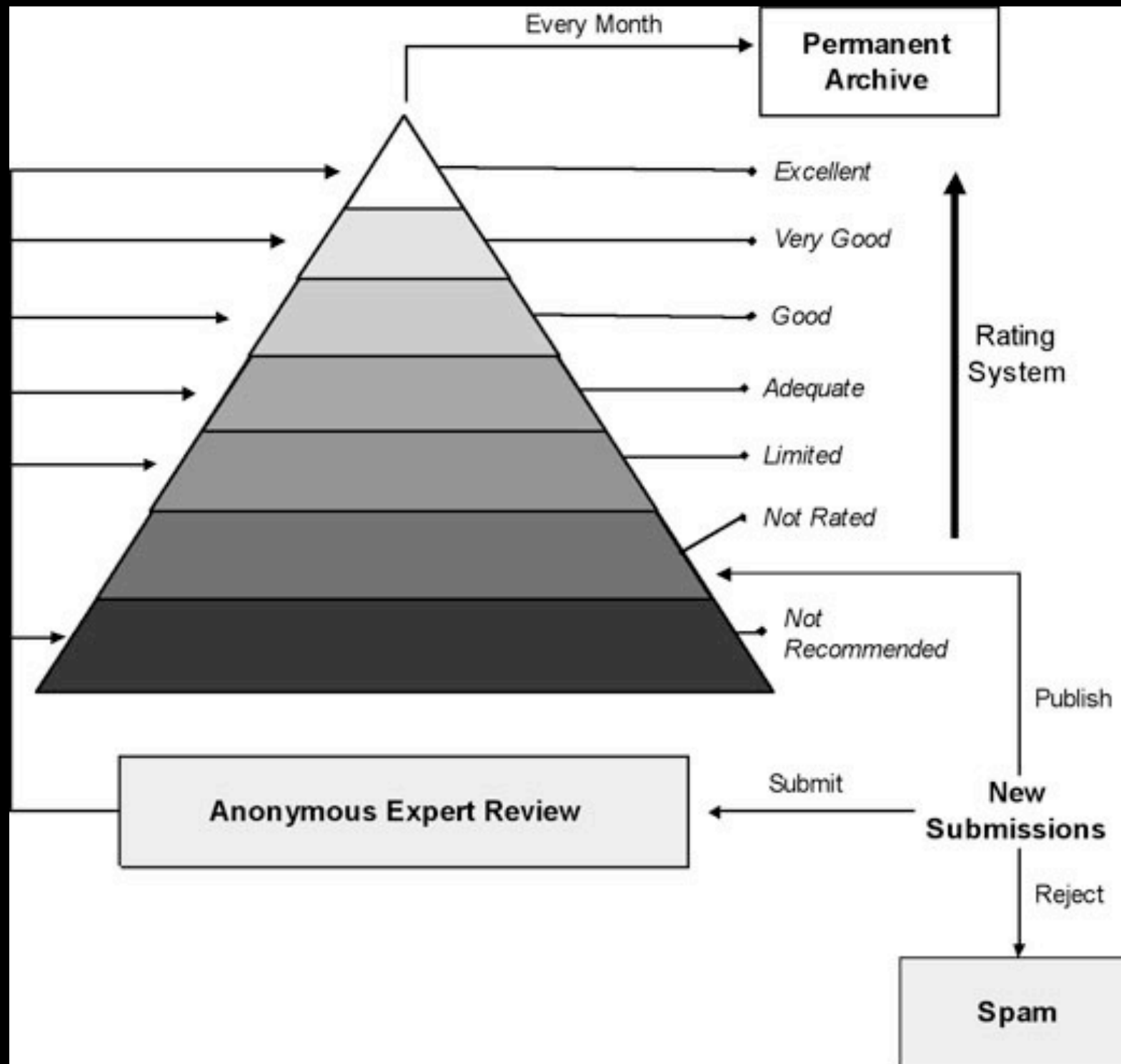
de processi

## **Accesso**

a documenti e a dati



# Un modello di pubblicazione aperto



[Whitman e  
Friedman]

# I più importanti archivi OA

1) Quello della propria istituzione

esempi:

Unitn - <http://eprints.biblio.unitn.it/>

LUISearch- <http://eprints.luiss.it/>

1b) In alternativa (prossimo futuro: OpenAIRE, <http://www.openaire.eu>)

2) Alcuni tra più noti archivi disciplinari:

ArXiv <http://arxiv.org>

RepeC-Research Papers in Economics <http://repec.org/>

Social Sciences Research Network - <http://www.ssrn.com/>

E-LIS - <http://eprints.rclis.org/>

NB: Il concetto di disciplina dev'essere ripensato. Esempi?

ArXiv, E-Lis, etc.

# Trovare gli archivi adeguati

**ROAR- Registry of Open Access Repositories**

<http://roar.eprints.org/>

**OpenDOAR - Directory of Open Access Repositories**

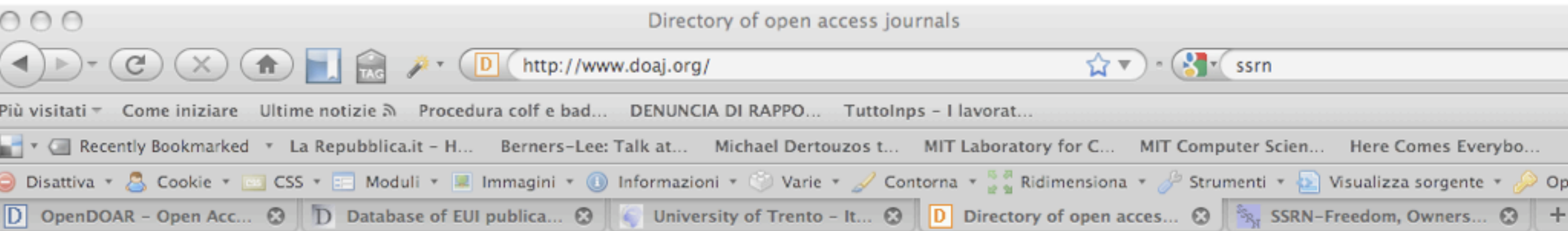
<http://www.opendoar.org/>



# DOAJ: riviste OA

**<http://www.doaj.org>**

- Lists free, full text, quality controlled scientific and scholarly journals in all subjects and languages.
- Currently 4833 journals are listed.
- Currently 1970 journals are searchable at article level.



**DOAJ** DIRECTORY OF  
OPEN ACCESS  
JOURNALS

**SPARC  
EUROPE  
AWARD  
2009**

For Outstanding  
Achievements  
in Scholarly  
Communications

**SPARC**  
The Scholarly Publishing and Academic Resources Coalition

[Find Journals](#)

[New titles](#)

[Find articles](#)

[Suggest a journal](#)

[About](#)

[FAQ](#)

[News](#)

Welcome to the Directory of Open Access Journals. This service covers free, full text, quality controlled scientific and scholarly journals. We aim to cover all subjects and languages. There are now **4844** journals in the directory. Currently **1970** journals are searchable at article level. As of today **377225** articles are included in the DOAJ service.

**News!** 5 February 2010, Stockholm & Lund: Co-Action Publishing and Lund University Libraries Head Office are pleased to announce the launch of the **Online Guide to Open Access Journals Publishing** at <http://www.doaj.org/bpguide>.

**Support the development and operation of DOAJ. Sign up for membership - go to the [membership page](#).**  
**We are very thankful for the support from those of you who have already decided to become DOAJ members. See the [list of members](#)**

# Altri strumenti per trovare letteratura OA (al di là di Google, GoogleScholar)

**Service Provider: Oaister - <http://oaister.worldcat.org/>**

- Contains records of digital resources from openarchive collections worldwide.
- More than 23 million records representing digital resources from more than 1,100 repositories
- not all records are accessible full text

**Where to go to search for OA content**

1) Scientific Commons, <http://www.scientificcommons.org>

- Contains records of digital resources from open-archive collections worldwide.
- More than 34 million records representing digital resources from more than 1,191 repositories
- not all records are accessible full text

**Driver - Digital Repository Infrastructure Vision for European Research**

<http://search.driver.research-infrastructures.eu/>

- Search for more than 21.700.00 documents in 252 repositories from 33 countries, mainly European



# In Italia

Offre accesso centralizzato alla letteratura scientifica archiviata nei repository italiani (over 23000 items available in full text format).

**PLEIADI-** Portale per la Letteratura scientifica Elettronica Italiana su Archivi aperti e Depositi

Istituzionali <http://www.openarchives.it/pleiadi/>

The screenshot shows a web browser window displaying the PLEIADI website. The browser's address bar shows the URL <http://www.openarchives.it/pleiadi/>. The website's header features the PLEIADI logo, which is a stylized constellation of stars. Below the header, the main navigation bar includes links for 'Login', 'Ricerca in OpenArchives', and 'News Repositories'. The 'Login' section on the left has fields for 'Nome utente:' and 'Password:', along with a 'Login utente' button. The 'Ricerca in OpenArchives' section in the center has a search input field with a 'Cerca' button, and it displays 'Record attualmente in archivio: 23879'. The 'News Repositories' section on the right features a 'Repository Platform News Feed' with a news item titled 'Fedora: Authentication Challenge Control - A Thought Experiment'. The left sidebar contains a 'Menu Principale' with links to 'Home', 'Ultime notizie', 'RSS: Notizie dal mondo OA', and 'Community & wiki'. The bottom of the page shows a list of recent news items under the heading 'Notizie recenti'.

**PLEIADI: Portale per la Letteratura scientifica Elettronica Italiana su Archivi aperti e Depositi Istituzionali**

**Login**

Nome utente:  
Password:  
[Login utente](#)

Hai perso la password?

**Registrati ora!**

**Menu Principale**

- Home
- Ultime notizie
- RSS: Notizie dal mondo OA
- Community & wiki

**Ricerca in OpenArchives**

[Cerca](#)

**Altre modalità di ricerca**  
Record attualmente in archivio: **23879**

**Notizie recenti**

- **Open come opportunità - Seminario CASPUR/CILEA** (26/2/2010)
- **Open Access 19 -23 ottobre - Materiali da utilizzare** (11/8/2009)
- **Pubblicati sul sito della CRUI: linee guida per gli archivi istituzionali e Open Access e la valutazione dei prodotti della ricerca** (5/8/2009)
- **Nuovo numero di D-lib Magazine** (19/5/2009)
- **Un articolo sull'atteggiamento dei bibliotecari accademici verso l'OA** (6/5/2009)

**News Repositories**

**Repository Platform News Feed**

**Fedora: Authentication Challenge Control - A Thought Experiment**

News Item edited by Chris Wilper  
Fedora's SOAP API has historically been split into two: API-A and API-M. The REST API, being resource-oriented, naturally isn't. And the control over which endpoints can be set to require authentication is inconsistent across the two. The old "Authenticate for API-A? Authenticate for API-M?" options no longer make sense when you look at Fedora's APIs as a whole. Here's one way the



# Gli editori e l'OA -

## SHERPA/RoMEO

<http://www.sherpa.ac.uk/romeo/>

### RoMEO colours

We have used different colours to help highlight publisher's archiving policies. These colours are a development from the original RoMEO project list, and differentiate between four categories of archiving rights:

ROMEO colour	Archiving policy
<a href="#">green</a>	can archive pre-print <i>and</i> post-print or publisher's version/PDF
<a href="#">blue</a>	can archive post-print (ie final draft post-refereeing) or publisher's version/PDF
<a href="#">yellow</a>	can archive pre-print (ie pre-refereeing)
<a href="#">white</a>	archiving not formally supported

Each publisher's entry is coded according to one of these colour categories.

The entry for each publisher also lists conditions or restrictions imposed by the publisher which govern archiving rights or activities. *Conditions* are taken as terms which can be easily accommodated and which do not hinder an author in archiving their work. A typical condition is to acknowledge the publisher's copyright in the work. *Restrictions* are more prohibitive, typically requiring some additional action on behalf of the author. Where a *Restriction* effectively blocks access to the eprint, such as in the case of an embargo on its public release, or requiring password-controlled access, then the partial archiving right is noted but the full colour categorisation does not apply.

Sometimes open access discussions talk about "gold" publishers. This is a later development independent of RoMEO categories, and is used to describe publishers of open access journals. For the purposes of archiving, all open access journals allow archiving and can be taken as RoMEO "green".

Some of the larger publishers have different archiving rights for different journals. This is particularly the case where they publish learned society journals on behalf of the society. A learned society might insist on a more liberal, or more restrictive archiving policy than the general publisher's copyright agreement allows. The RoMEO colour coding relates to the overall permissions given by a publisher. For example, a publisher has to apply

# Policies e Mandati OA: UE

**European Research Council** has adopted a mandatory policy - EU funded research (6 months embargo)

**EU – FP7 (2007-2013) Pilot Project on Open Access (2008-2013)**

7 disciplines (Energy, Environment, Health, ICT, E-Infrastructures, Science in Society, Socio-economic Sciences and Humanities)

- Project OpenAire

**European Science Foundation (ESF) e EuroHORCS (European Head of Research Council):** recommend open access to all public research funding agencies in Europe and are about to issue mandatory policies

# Linee guida della CRUI

<http://www.cruir.it/HomePage.aspx?ref=116>

## **Raccomandazioni alle università:**

- archivi istituzionali
- tesi di dottorato
- riviste
- valutazione della ricerca



# Raccomandazioni CRUI / Archivi istituzionali

<http://www.cruai.it/HomePage.aspx?ref=1781>

## **Adeguarsi alle linee-guida ERC (2007):**

tutte le pubblicazioni validate da peer review derivanti da progetti di ricerca finanziati con fondi europei dovranno essere rese accessibili in un archivio istituzionale o disciplinare appropriato, entro 6 mesi dalla pubblicazione.

Politiche analoghe sono adottate o in corso di adozione da parte di stati e enti di ricerca (**NIH, Wellcome Trust, Max Planck, VII programma quadro EU, Harvard, Stanford, valutazione della ricerca britannica e australiana...**).

**Come riempirli?** Integrazione/interoperabilità con: didattica, anagrafe della ricerca, university press, riviste, web semantico...

# Raccomandazioni CRUI / Tesi di dottorato

<http://www.cruui.it/HomePage.aspx?ref=1149>

- Il **deposito legale delle tesi** di dottorato presso le Biblioteche Nazionali di Roma e Firenze è un obbligo.
- La volontà e i diritti dei dottorandi/dottori di ricerca non possono eludere quest'obbligo di legge.
- La volontà e i diritti dei dottorandi/dottori di ricerca non possono neppure compromettere le libere scelte dei soggetti pubblici "Università" sull'Open Access.
- Il deposito legale delle tesi di dottorato comporta che esse possano essere oggetto di "servizi bibliografici nazionali di informazione e accesso".
- La consultabilità delle tesi di dottorato per i terzi è un diritto che può avere alcune limitazioni circoscritte, temporanee, motivate: sempre legate alla disciplina sulla tutela della proprietà intellettuale e limitatamente ai casi da essa previsti.
- **I dottorandi mantengono integri il diritto d'autore sul testo della tesi:** per il fatto stesso di partecipare ad un corso che si conclude con un testo scritto, confermano e vedono resi espliciti a data certa i propri diritti

# Raccomandazioni CRUI / Riviste

<http://www.cruir.it/HomePage.aspx?ref=1789>

**Serial price crisis:** il mercato oligopolistico dei core journal ha portato a un aumento del 200% oltre l'inflazione (1986-2004) delle tariffe di abbonamento.

Le università, che offrono gratis gli articoli agli editori, li devono ricomprare a caro prezzo, una volta incorporati nelle riviste, e gravati di vincoli.

Il modello readers pay, oltre a costringere le università a pagare per accedere a quello che esse stesse hanno prodotto, fa circolare poco i testi e penalizza le istituzioni povere. Il modello authors pay, diffuso nel mondo umanistico, oltre a penalizzare le istituzioni povere, non garantisce, di per sé, l'accesso aperto.

## **Soluzione: institutions pay:**

- invece di regalare i propri prodotti agli editori per doverli ricomprare a caro prezzo, finanziare la pubblicazione ad accesso aperto e il referaggio.
- avviare forme di sperimentazione parallele alla peer review tradizionale, quali la soft peer review, la peer review ex post e la selezione attraverso l'uso.



# Raccomandazioni CRUI / Valutazione della ricerca

<http://www.cruai.it/HomePage.aspx?ref=1782>

- **anagrafe della ricerca interoperabile con gli standard minimi OA e OAI** e con la “Direttiva Stanca” sull’Open Source e la “Legge Stanca” sull’accessibilità;
- valutazione dei prodotti della ricerca secondo criteri sia qualitativi (peerreview) che quantitativi, utilizzando anche indicatori biblio- e webometrici alternativi all’Impact Factor e alla Citation Analysis
- considerare a fini di valutazione non soltanto la cosiddetta “versione editoriale”, ma anche le versioni “pre-print dell’autore” o “post-print referato dell’autore”, qualora scientificamente rilevanti;
- estendere la valutazione ai prodotti di ricerca pubblicati su riviste ad accesso aperto peer-reviewed.

# 2012

- > **270 istituzioni** di ricerca hanno sottoscritto la Dichiarazione di Berlino del 2003

- > **170 policies** e mandati registrati, tra cui  
**Politiche dei finanziatori pubblici: UE**

- FP7 Pilot

- European Research Council

**Esempi virtuosi:**

- JISC

- Agenzie della valutazione UK e Australia

- ERC?

**Strumenti:**

- es. Sherpa/ROMeO

- > **1500 archivi** OAI-PMH compliant

(cfr. OpenDOAR, ROAR)

- > **4500 peer reviewed journals** sono ad accesso aperto

(cfr. DOAJ)

# in Italia

- **74 università** su 77 hanno sottoscritto la Dichiarazione di Berlino
- **CRUI**: linee guida 2009
- **Gli archivi istituzionali OA** sono 28
- **Policy dei finanziatori e delle istituzioni:**  
es. Trento: integrazione archivio istituzionale e anagrafe della ricerca.  
es. Tuscia: incentivi in sede di valutazione



Copyright  
(o meglio: diritto d'autore)  
e licenze

# Il diritto d'autore (in Italia)

Legge 22 aprile 1941, n. 633 (LDA)

Il diritto d'autore dura 70 anni dopo la morte dell'autore.

E' composto da

- **diritto morale** (inalienabile)
- **diritti economici** (alienabili)

Il testo della legge: <[http://www.interlex.it/testi/l41\\_633.htm](http://www.interlex.it/testi/l41_633.htm)>

Per saperne di più:

A. De Robbio, La normativa italiana sul diritto d'autore

<<http://www.math.unipd.it/~derobbio/dd/copyr06.htm>>

# Un'alternativa (deroga) al diritto d'autore: le licenze

Tramite l'applicazione di licenze, l'autore può decidere di cedere in tutto o in parte i suoi **diritti economici** sull'opera.

Da dove nascono le licenze?

# Regime giuridico del software

## Software proprietario

Viene rilasciato solo il codice oggetto. Acquistandolo, si può usare ma non si può “leggere”.

## Software FLOSS

Vengono rilasciati codice sorgente e codice oggetto.

## Software Libero

Vengono rilasciati codice sorgente e codice oggetto a particolari **condizioni**. Quali?





# Software libero

"libertà 0". Libertà di eseguire il programma per qualsiasi scopo

"libertà 1". Libertà di studiare il programma e modificarlo

"libertà 2". Libertà di copiare il programma in modo da aiutare il prossimo

"libertà 3". Libertà di migliorare il programma e di distribuirne pubblicamente i miglioramenti, in modo tale che tutta la comunità ne tragga beneficio

NB. Poiché "free" si riferisce alla libertà e non al prezzo, vendere copie di un programma non contraddice il concetto di software libero. In effetti, la libertà di vendere copie di programmi è essenziale.



*Richard Stallman*

# Software libero: storia



- **1983** inizia a lavorare al progetto **GNU** per dedicarsi a tempo pieno; abbandona definitivamente il MIT nel gennaio del 1984, per evitarne le possibili influenze;
- elabora il **Manifesto GNU** e nel **1985** fonda la **Free Software Foundation**.
- **1989** inventa il concetto di **copyleft**.



## **Copyleft:** permesso d'autore

Il permesso d'autore (copyleft) usa le leggi sul diritto d'autore (copyright), ma le capovolge per ottenere lo scopo opposto: invece che un metodo per privatizzare il software, diventa infatti un mezzo per mantenerlo libero.

Il succo dell'idea di permesso d'autore consiste nel dare a chiunque il permesso di eseguire il programma, copiare il programma, modificare il programma, e distribuirne versioni modificate, ma senza dare il permesso di aggiungere restrizioni.

In tal modo, le libertà essenziali che definiscono il “free software” (software libero) sono garantite a chiunque ne abbia una copia, e diventano diritti inalienabili.



# GNU GPL: General Public License

Una licenza copyleft che garantisce le 4 libertà.



# Copyleft e contenuti

L'autore cede parte dei suoi diritti economici al lettore. Un'applicazione pratica: le licenze Creative Commons.

**Le licenze creative commons sono nate per testi, musica e video, e permettono:**

- a) di ottenere senza permesso opere derivate e/o
- b) di vendere l'opera senza permesso

**NB: Il diritto morale dell'autore (paternità intellettuale) in Europa è inalienabile.**

# Bibliografia

P. Gargiulo, *Open Access*, presentazione online all'URL:  
[http://bfp.sp.unipi.it/~didonato/ovre/spring\\_school/gargiulo.pdf](http://bfp.sp.unipi.it/~didonato/ovre/spring_school/gargiulo.pdf)

J.-C. Guédon, *Open Access. Contro gli oligopoli nel sapere*, ETS, Pisa 2009

<http://www.edizioniets.com/Scheda.asp?N=9788846725172>

M.C. Pievatolo, *Per una ragione pubblica: le linee guida della CRUI per l'accesso aperto*, OADay 2009, Pisa, ottobre 2009,  
<http://eprints.rclis.org/17014/>